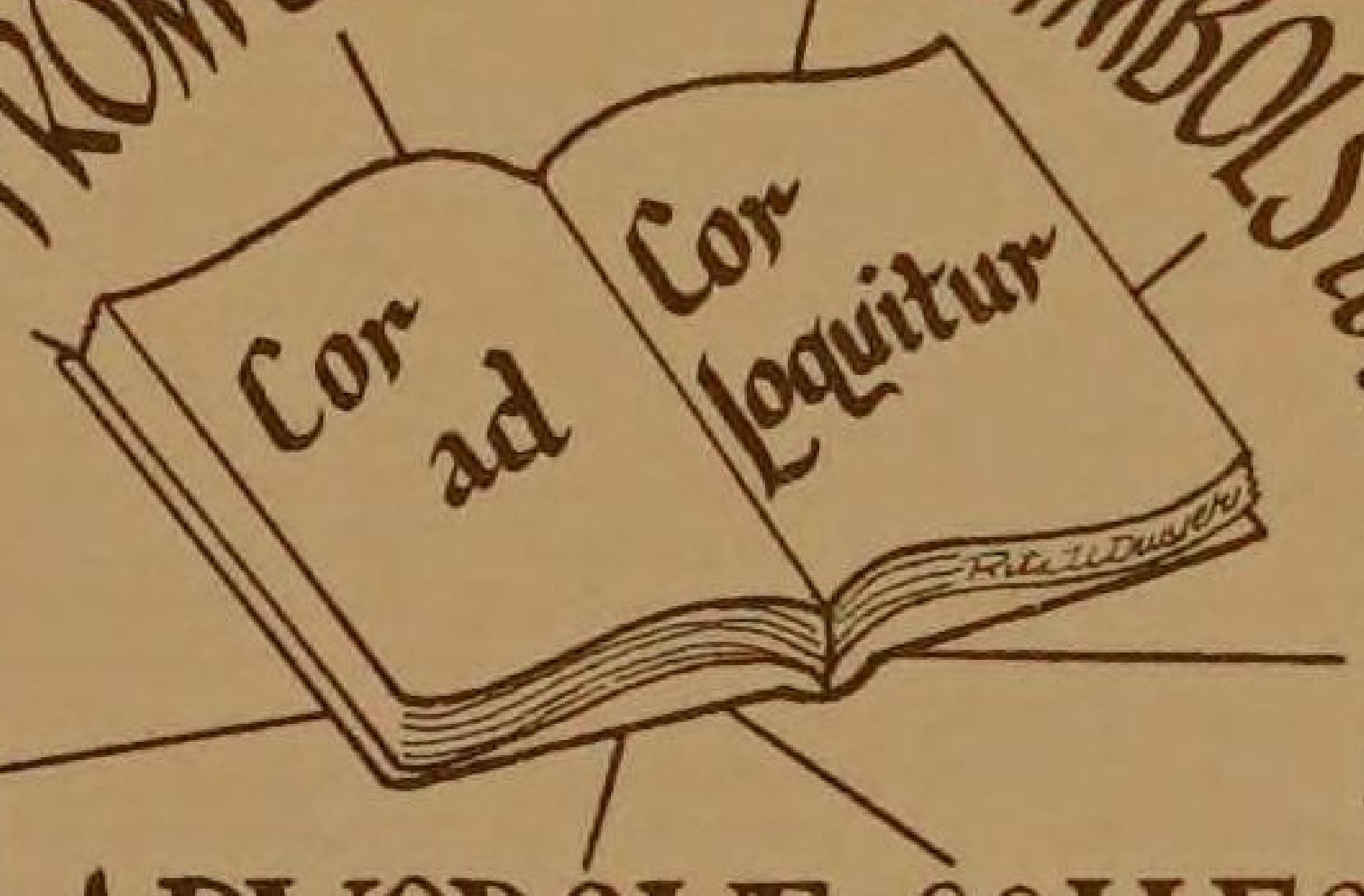


FROM SHADOWS and SYMBOLS to the TRUTH

MARYGROVE COLLEGE
EX LIBRIS

530.12
F43

Quantum Mechanics
and Path Integrals

Quantum Mechanics
and Path Integrals

INTERNATIONAL SERIES IN PURE AND APPLIED PHYSICS

Leonard I. Schiff, *Consulting Editor*

Adler, Bazin, and Schiffer Introduction to General Relativity
Allis and Herlin Thermodynamics and Statistical Mechanics
Becker Introduction to Theoretical Mechanics
Bjorken and Drell Relativistic Quantum Mechanics
Bjorken and Drell Relativistic Quantum Fields
Chodorow and Susskind Fundamentals of Microwave Electronics
Clark Applied X-rays
Collin Field Theory of Guided Waves
Evans The Atomic Nucleus
Feynman and Hibbs Quantum Mechanics and Path Integrals
Ginzton Microwave Measurements
Green Nuclear Physics
Gurney Introduction to Statistical Mechanics
Hall Introduction to Electron Microscopy
Hardy and Perrin The Principles of Optics
Harnwell Electricity and Electromagnetism
Harnwell and Livingood Experimental Atomic Physics
Henley and Thirring Elementary Quantum Field Theory
Houston Principles of Mathematical Physics
Kennard Kinetic Theory of Gases
Lane Superfield Physics
Leighton Principles of Modern Physics
Lindsay Mechanical Radiation
Livingston and Blewett Particle Accelerators
Middleton An Introduction to Statistical Communication Theory

Morse Vibration and Sound
Morse and Feshbach Methods of Theoretical Physics
Muskat Physical Principles of Oil Production
Present Kinetic Theory of Gases
Read Dislocations in Crystals
Richtmyer, Kennard, and Lauritsen Introduction to Modern Physics
Schiff Quantum Mechanics
Seitz The Modern Theory of Solids
Slater Introduction to Chemical Physics
Slater Quantum Theory of Atomic Structure, Vol. I
Slater Quantum Theory of Atomic Structure, Vol. II
Slater Quantum Theory of Matter
Slater Quantum Theory of Molecules and Solids, Vol. 1
Slater Quantum Theory of Molecules and Solids, Vol. 2
Slater and Frank Electromagnetism
Slater and Frank Introduction to Theoretical Physics
Slater and Frank Mechanics
Smythe Static and Dynamic Electricity
Stratton Electromagnetic Theory
Tinkham Group Theory and Quantum Mechanics
Townes and Schawlow Microwave Spectroscopy
White Introduction to Atomic Spectra

The late F. K. Richtmyer was Consulting Editor of the series from its inception in 1929 to his death in 1939. Lee A. DuBridge was Consulting Editor from 1939 to 1946; and G. P. Harnwell from 1947 to 1954.

R. P. Feynman

Tolman Professor of Physics
California Institute of Technology

A. R. Hibbs

Jet Propulsion Laboratory
California Institute of Technology

McGraw-Hill Book Company

New York St. Louis San Francisco Toronto London Sydney

Quantum
Mechanics
and
Path Integrals

QUANTUM MECHANICS AND PATH INTEGRALS

Copyright © 1965 by McGraw-Hill, Inc. All Rights Reserved.
Printed in the United States of America. This book, or
parts thereof, may not be reproduced in any form without
permission of the publishers.

Library of Congress Catalog Card Number 64-25171

20650

.234567890 MP 721069876

Preface

The fundamental physical and mathematical concepts which underlie the path integral approach to quantum mechanics were first developed by R. P. Feynman in the course of his graduate studies at Princeton, although more fully developed ideas, such as those described in this volume, were not worked out until a few years later. These early inquiries were involved with the problem of the infinite self-energy of the electron. In working on that problem, a "least-action" principle using half advanced and half retarded potentials was discovered. The principle could deal successfully with the infinity arising in the application of classical electrodynamics.

The problem then became one of applying this action principle to quantum mechanics in such a way that classical mechanics could arise naturally as a special case of quantum mechanics when h was allowed to go to zero.

Feynman searched for any ideas which might have been previously worked out in connecting quantum-mechanical behavior with such classical ideas as the lagrangian or, in particular, Hamilton's principle function S, the indefinite integral of the lagrangian. During some conversations with a visiting European physicist, Feynman learned of a paper in which Dirac had suggested that the exponential function of $i\epsilon$ times the lagrangian was analogous to a transformation function for the quantum-mechanical wave function in that the wave function at one moment could be related to the wave function at the next moment (a time interval ϵ later) by multiplying with such an exponential function.

The question that then arose was what Dirac had meant by the phrase "analogous to," and Feynman determined to find out whether or not it would be possible to substitute the phrase "equal to." A brief analysis showed that indeed this exponential function could be used in this manner directly.

Further analysis then led to the use of the exponent of the time integral of the lagrangian, S (in this volume referred to as the *action*) as the transformation function for finite time intervals. However, in the application of this function it is necessary to carry out integrals over all space variables at every instant of time.

In preparing an article[1] describing this idea, the idea of "integral over all paths" was developed as a way of both describing and evaluating the required integrations over space coordinates. By this time a number of mathematical devices had been developed for applying the path integral technique and a number of special applications had been worked out, although the primary direction of work at this time was toward quantum electrodynamics. Actually, the path integral did not then provide, nor has it since provided, a truly satisfactory method of avoiding the divergence difficulties of quantum electrodynamics, but it has been found to be most useful in solving other problems in that field. In particular, it provides an expression for quantum-electrodynamic laws in a form that relativistic invariance is obvious. In addition, useful applications to other problems of quantum mechanics have been found.

The most dramatic early application of the path integral method to an intractable quantum-mechanical problem followed shortly after the discovery of the Lamb shift and the subsequent theoretical diffi-

[1] R. P. Feynman, Space-Time Approach to Non-relativistic Quantum Mechanics, *Rev. Mod. Phys.*, vol. 20, p. 367, 1948.

culties in explaining this shift without obviously artificial means of getting rid of divergent integrals. The path integral approach provided one way of handling these awkward infinities in a logical and consistent manner.

The path integral approach was used as a technique for teaching quantum mechanics for a few years at the California Institute of Technology. It was during this period that A. R. Hibbs, a student of Feynman's, began to develop a set of notes suitable for converting a lecture course on the path integral approach to quantum mechanics into a book on the same subject.

Over the succeeding years, as the book itself was elaborated, other subjects were brought into both the lectures of Dr. Feynman and the book; examples are statistical mechanics and the variational principle. At the same time, Dr. Feynman's approach to teaching the subject of quantum mechanics evolved somewhat away from the initial path integral approach. At the present time, it appears that the operator technique is both deeper and more powerful for the solution of more general quantum-mechanical problems. Nevertheless, the path integral approach provides an intuitive appreciation of quantum-mechanical behavior which is extremely valuable in gaining an intuitive appreciation of quantum-mechanical laws. For this reason, in those fields of quantum mechanics where the path integral approach turns out to be particularly useful, most of which are described in this book, the physics student is provided with an excellent grasp of basic quantum-mechanical principles which will permit him to be more effective in solving problems in broader areas of theoretical physics.

R. P. Feynman

A. R. Hibbs

Contents

Preface *vii*

chapter 1 **The Fundamental Concepts of Quantum Mechanics** *1*

1-1 Probability in Quantum Mechanics 2
1-2 The Uncertainty Principle 9
1-3 Interfering Alternatives 13
1-4 Summary of Probability Concepts 19
1-5 Some Remaining Thoughts 22
1-6 The Purpose of This Book 23

chapter 2 **The Quantum-mechanical Law of Motion** *25*

2-1 The Classical Action 26
2-2 The Quantum-mechanical Amplitude 28
2-3 The Classical Limit 29
2-4 The Sum over Paths 31
2-5 Events Occurring in Succession 36
2-6 Some Remarks 38

chapter 3 **Developing the Concepts with Special Examples** *41*

3-1 The Free Particle 42
3-2 Diffraction through a Slit 47
3-3 Results for a Sharp-edged Slit 55
3-4 The Wave Function 57
3-5 Gaussian Integrals 58
3-6 Motion in a Potential Field 62
3-7 Systems with Many Variables 65
3-8 Separable Systems 66
3-9 The Path Integral as a Functional 68
3-10 Interaction of a Particle and a Harmonic Oscillator 69
3-11 Evaluation of Path Integrals by Fourier Series 71

chapter 4 **The Schrödinger Description of Quantum Mechanics** *75*

4-1 The Schrödinger Equation 76
4-2 The Time-independent Hamiltonian 84
4-3 Normalizing the Free-particle Wave Functions 89

chapter 5 **Measurements and Operators** *95*

5-1 The Momentum Representation 96
5-2 Measurement of Quantum-mechanical Variables 106
5-3 Operators 112

chapter 6 **The Perturbation Method in Quantum Mechanics** *119*

6-1 The Perturbation Expansion 120
6-2 An Integral Equation for K_V 126
6-3 An Expansion for the Wave Function 127
6-4 The Scattering of an Electron by an Atom 129
6-5 Time-dependent Perturbations and Transition Amplitudes 143

chapter 7 **Transition Elements** *163*

7-1 Definition of the Transition Element 164
7-2 Functional Derivatives 170
7-3 Transition Elements of Some Special Functionals 174
7-4 General Result for Quadratic Actions 182
7-5 Transition Elements and the Operator Notation 184
7-6 The Perturbation Series for a Vector Potential 189
7-7 The Hamiltonian 192

chapter 8 **Harmonic Oscillators** *197*

8-1 The Simple Harmonic Oscillator 198
8-2 The Polyatomic Molecule 203
8-3 Normal Coordinates 208
8-4 The One-dimensional Crystal 212
8-5 The Approximation of Continuity 218
8-6 Quantum Mechanics of a Line of Atoms 222
8-7 The Three-dimensional Crystal 224
8-8 Quantum Field Theory 229
8-9 The Forced Harmonic Oscillator 232

chapter 9 **Quantum Electrodynamics** *235*

9-1 Classical Electrodynamics 237
9-2 The Quantum Mechanics of the Radiation Field 242
9-3 The Ground State 244
9-4 Interaction of Field and Matter 247
9-5 A Single Electron in a Radiative Field 253
9-6 The Lamb Shift 256
9-7 The Emission of Light 260
9-8 Summary 262

chapter 10 **Statistical Mechanics** *267*

10-1 The Partition Function 269
10-2 The Path Integral Evaluation 273
10-3 Quantum-mechanical Effects 279
10-4 Systems of Several Variables 287
10-5 Remarks on Methods of Derivation 295

chapter 11 **The Variational Method** *299*

11-1 A Minimum Principle 300
11-2 An Application of the Variational Method 303
11-3 The Standard Variational Principle 307
11-4 Slow Electrons in a Polar Crystal 310

chapter 12 **Other Problems in Probability** *321*

12-1 Random Pulses 322
12-2 Characteristic Functions 324
12-3 Noise 326
12-4 Gaussian Noise 331
12-5 Noise Spectrum 334
12-6 Brownian Motion 336
12-7 Quantum Mechanics 340
12-8 Influence Functionals 343
12-9 Influence Functional from a Harmonic Oscillator 351
12-10 Conclusions 355

Appendix **Some Useful Definite Integrals** *357*

Index *359*

1

The Fundamental Concepts of Quantum Mechanics

1-1 PROBABILITY IN QUANTUM MECHANICS[1]

From about the beginning of the twentieth century experimental physics amassed an impressive array of strange phenomena which demonstrated the inadequacy of classical physics. The attempts to discover a theoretical structure for the new phenomena led at first to a confusion in which it appeared that light, and electrons, behaved sometimes like waves and sometimes like particles. This apparent inconsistency was completely resolved in 1926 and 1927 in the theory called quantum mechanics. The new theory asserts that there are experiments for which the exact outcome is fundamentally unpredictable and that in these cases one has to be satisfied with computing probabilities of various outcomes. But far more fundamental was the discovery that in nature the laws of combining probabilities were *not* those of the classical probability theory of Laplace. The quantum-mechanical laws of the physical world approach very closely the laws of Laplace as the size of the objects involved in the experiments increases. Therefore, the laws of probabilities which are conventionally applied are quite satisfactory in analyzing the behavior of the roulette wheel but not the behavior of a single electron or a photon of light.

A Conceptual Experiment. The concept of probability is not altered in quantum mechanics. When we say the probability of a certain outcome of an experiment is p, we mean the conventional thing, i.e., that if the experiment is repeated many times, one expects that the fraction of those which give the outcome in question is roughly p. We shall not be at all concerned with analyzing or defining this concept in more detail; for no departure from the concept used in classical statistics is required.

What is changed, and changed radically, is the method of calculating probabilities. The effect of this change is greatest when dealing with objects of atomic dimensions. For this reason we shall illustrate the laws of quantum mechanics by describing the results to be expected in some conceptual experiments dealing with a single electron.

Our imaginary experiment is illustrated in Fig. 1-1. At A we have a source of electrons S. The electrons at S all have the same energy but come out in all directions to impinge on a screen B. The screen

---

[1] Much of the material appearing in this chapter was originally presented as a lecture by R. P. Feynman and published as The Concept of Probability in Quantum Mechanics in the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, Calif., 1951.

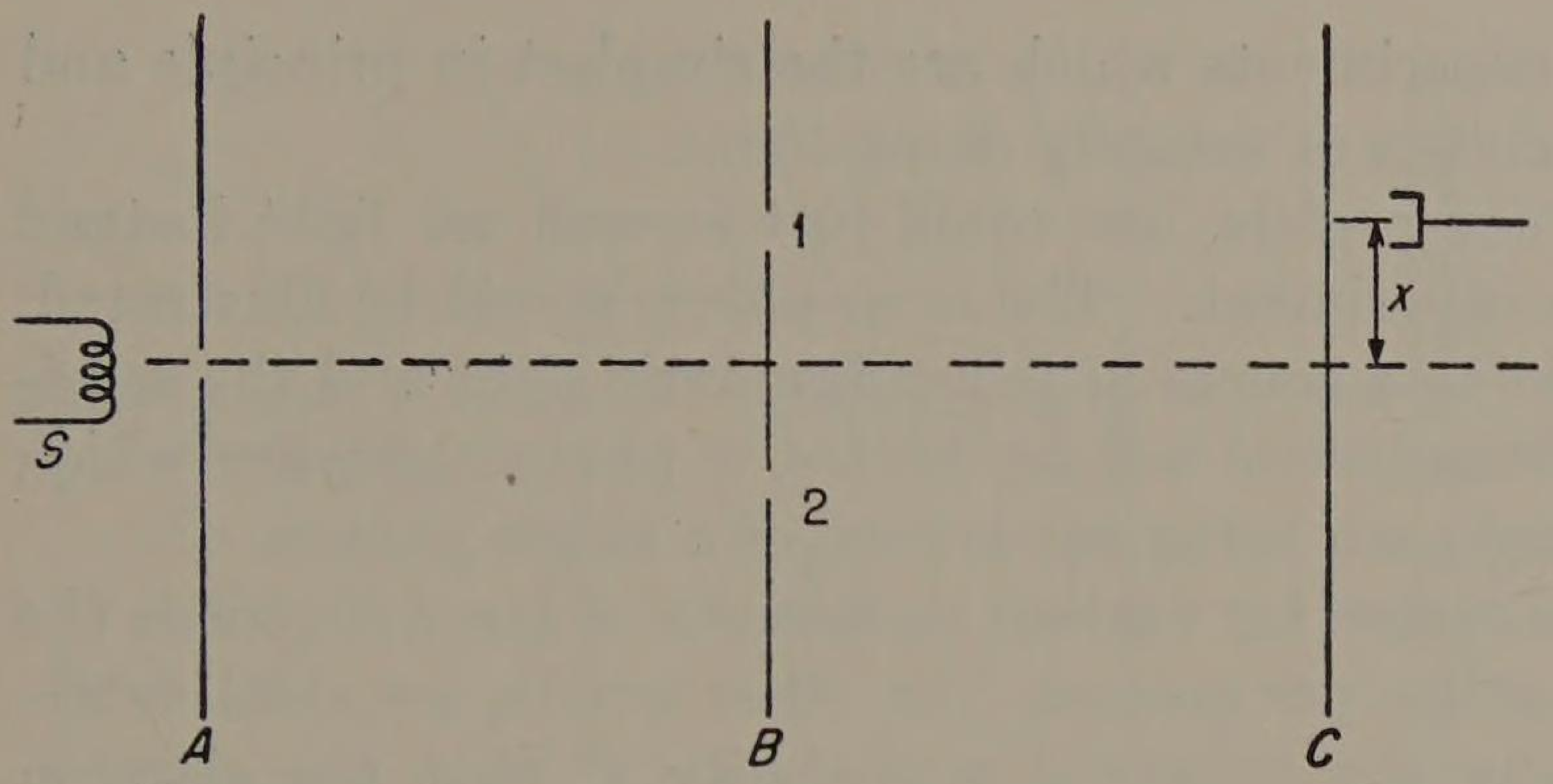


Fig. 1-1 The experimental arrangement. Electrons emitted at A make their way to the detector at screen C, but a screen B with two holes is interposed. The detector registers a count for each electron which arrives; the fraction which arrives when the detector is placed at a distance x from the center of the screen is measured and plotted against x, as in Fig. 1-2.

B has two holes, 1 and 2, through which the electrons may pass. Finally, behind the screen B at a plane C we have a detector of electrons which may be placed at various distances x from the center of the screen.

If the detector is extremely sensitive (as a Geiger counter is) it will be discovered that the current arriving at x is not continuous, but corresponds to a rain of particles. If the intensity of the source S is very low, the detector will record pulses representing the arrival of individual particles, separated by gaps in time during which nothing arrives. This is the reason we say electrons are particles. If we had detectors simultaneously all over the screen C, with a very weak source S, only one detector would respond, then after a little time, another would record the arrival of an electron, etc. There would never be a half response of the detector; either an entire electron would arrive or nothing would happen. And two detectors would never respond simultaneously (except for the coincidence that the source emitted two electrons within the resolving time of the detectors—a coincidence whose probability can be decreased by further decrease of the source intensity). In other words, the detector of Fig. 1-1 records the passage of a single corpuscular entity traveling from S through a hole in screen B to the point x.

This particular experiment has never been done in just this way. In the following description we are stating what the results would be according to the laws which fit every experiment of this type which has ever been performed. Some experiments which directly illustrate the conclusions we are reaching here have been done, but such experiments are usually more complicated. We prefer, for pedagogical reasons, to select experiments which are the simplest in principle and disregard the difficulties of actually doing them.

Incidentally, if one prefers, one could just as well use light instead of electrons in this experiment. The same points would be illustrated. The source S could be a source of monochromatic light and the sensitive detector a photoelectric cell or, better, a photomultiplier which would record pulses, each being the arrival of a single photon.

What we shall measure for various positions x of the detector is the mean number of pulses per second. In other words, we shall determine experimentally the (relative) probability P that the electron passes from S to x, as a function of x.

The graph of this probability as a function of x is the complicated curve illustrated qualitatively in Fig. 1-2*a*. It has several maxima and minima, and there are locations near the center of the screen at which electrons hardly ever arrive. It is the problem of physics to discover the laws governing the structure of this curve.

We might at first suppose (since the electrons behave as particles) that

I. Each electron which passes from S to x must go through either hole 1 or hole 2. As a consequence of I we expect that:

II. The chance of arrival at x should be the sum of two parts, P_1, the chance of arrival coming through hole 1, plus P_2, the chance of arrival coming through hole 2.

We may find out if this is true by direct experiment. Each of the component probabilities is easy to determine. We simply close hole

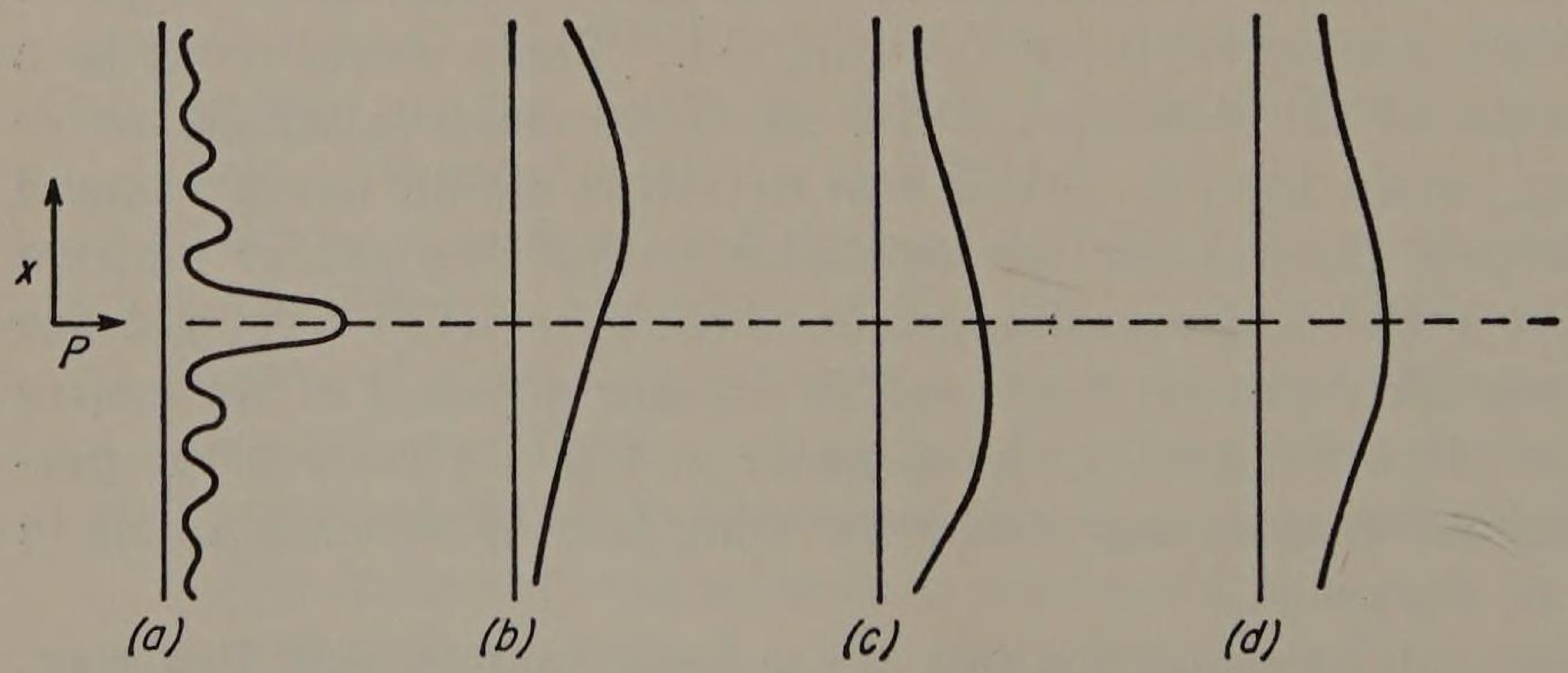


Fig. 1-2 Results of the experiment. Probability of arrival of electrons at x plotted against the position x of the detector. The result of the experiment of Fig. 1-1 is plotted here at (a). If only one hole is open, so the electrons can go through just hole 1, the result is (b). For just hole 2 open, it is (c). If we imagine each electron just goes through one hole or the other, we expect the curve (d) = (b) + (c) when both holes are open. This is considerably different from what we actually get, (a).

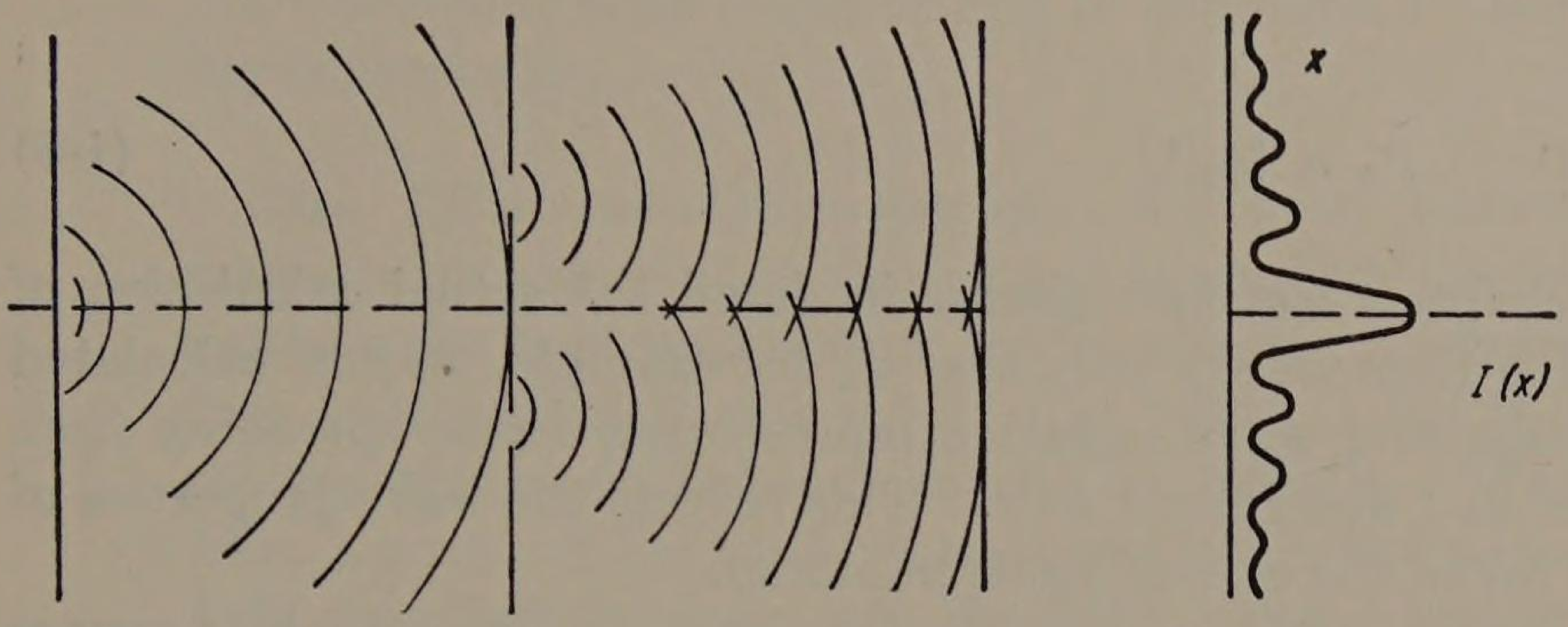


Fig. 1-3 An analogous experiment in wave interference. The complicated curve $P(x)$ in Fig. 1-2*a* is the same as the intensity $I(x)$ of waves which would arrive at x starting from S and coming through the holes. At some points x the wavelets from holes 1 and 2 interfere destructively (e.g., a crest from hole 1 arrives at the same time as a trough from hole 2); at others, constructively. This produces the complicated minima and maxima of the curve $I(x)$.

2 and measure the chance of arrival at x with only hole 1 open. This gives the chance P_1 of arrival at x for electrons coming through 1. The result is given in Fig. 1-2*b*. Similarly, by closing 1 we find the chance P_2 of arrival through hole 2 (Fig. 1-2*c*).

The sum of these (Fig. 1-2*d*) clearly does not agree with the curve (*a*). Hence, experiment tells us definitely that $P \neq P_1 + P_2$, or that II is false.

The Probability Amplitude. The chance of arrival at x with both holes open is *not* the sum of the chance with just hole 1 open plus that with just hole 2 open.

Actually, the complicated curve $P(x)$ is familiar, inasmuch as it is exactly the intensity of distribution in the interference pattern to be expected if waves starting from S pass through the two holes and impinge on the screen C (Fig. 1-3). The easiest way to represent wave amplitudes is by complex numbers. We can state the correct law for $P(x)$ mathematically by saying that $P(x)$ is the absolute square of a certain complex quantity (if electron spin is taken into account, it is a hypercomplex quantity) $\phi(x)$ which we call the *probability amplitude* of arrival at x. Furthermore, $\phi(x)$ is the sum of two contributions: ϕ_1, the amplitude of arrival through hole 1, plus ϕ_2, the amplitude of arrival through hole 2. In other words,

III. There are complex numbers ϕ_1 and ϕ_2 such that

$$P = |\phi|^2 \tag{1-1}$$

$$\phi = \phi_1 + \phi_2 \tag{1-2}$$

and

$$P_1 = |\phi_1|^2 \qquad P_2 = |\phi_2|^2 \tag{1-3}$$

In later chapters we shall discuss in detail the actual calculation of ϕ_1 and ϕ_2. Here we say only that ϕ_1, for example, may be calculated as a solution of a wave equation representing waves spreading from the source to 1 and from 1 to x. This reflects the wave properties of electrons (or in the case of light, photons).

To summarize: We *compute* the intensity (i.e., the absolute square of the amplitude) of waves which would arrive in the apparatus at x and then *interpret* this intensity as the probability that a particle will arrive at x.

Logical Difficulties. What is remarkable is that this dual use of wave and particle ideas does not lead to contradictions. This is so only if great care is taken as to what kind of statements one is permitted to make about the experimental situation.

To discuss this point in more detail, consider first the situation which arises from the observation that our new law III of composition of probabilities implies, in general, that it is not true that $P = P_1 + P_2$. We must conclude that when both holes are open, it is *not true* that the particle goes through one hole or the other. For if it had to go through one or the other, we could classify all the arrivals at x into two disjoint classes, namely, those arriving via hole 1 and those arriving through hole 2; and the frequency P of arrival at x would surely be the sum of the frequency P_1 of particles coming through hole 1 and the frequency P_2 of those coming through hole 2.

To extricate ourselves from the logical difficulties introduced by this startling conclusion, we might try various artifices. We might say, for example, that perhaps the electron travels in a complex trajectory going through hole 1, then back through hole 2 and finally out through 1 in some complicated manner. Or perhaps the electron spreads out somehow and passes partly through both holes so as to eventually produce the interference result III. Or perhaps the chance P_1 that the electron passes through hole 1 has not been determined correctly inasmuch as closing hole 2 might have influenced the motion near hole 1. Many such classical mechanisms have been tried to explain the result. When light photons are used (in which case the same law III applies), the two interfering paths 1 and 2 can be made to be many centimeters apart (in space), so that the two alternative trajectories must almost certainly be independent. That the actual situation is more profound than might at first be supposed is shown by the following experiment.

The Effect of Observation. We have concluded on logical grounds that since $P \neq P_1 + P_2$, it is not true that we can expect to analyze the electron's motion by the simple assumption that the electron passes through either hole 1 or hole 2. But it is easy to design an experiment to test our conclusion directly. We have merely to have a source of light behind the holes and watch to see through which hole the electron passes (see Fig. 1-4). For electrons scatter light, so that if light is scattered behind hole 1, we may conclude that an electron passed through hole 1; and if it is scattered in the neighborhood of hole 2, the electron has passed through hole 2.

The result of this experiment is to show unequivocally that the electron *does* pass through either hole 1 or hole 2! That is, for every electron which arrives at the screen C (assuming the light is strong enough that we do not miss seeing it) light is scattered either behind hole 1 or behind hole 2, and never (if the source S is very weak) at both places. (A more delicate experiment could even show that the charge passing through the holes passes through either one or the other and is in all cases the complete charge of one electron and not a fraction of it.)

It now appears that we have come to a paradox. For suppose that we combine the two experiments. We watch to see through which hole the electron passes and at the same time measure the chance that the electron arrives at x. Then for each electron which arrives at x we can say experimentally whether it came through hole 1 or hole 2. First we may verify that P_1 is given by curve (b), because if we select, of the electrons which arrive at x, only those which appear to come through hole 1 (by scattering light there), we find they are,

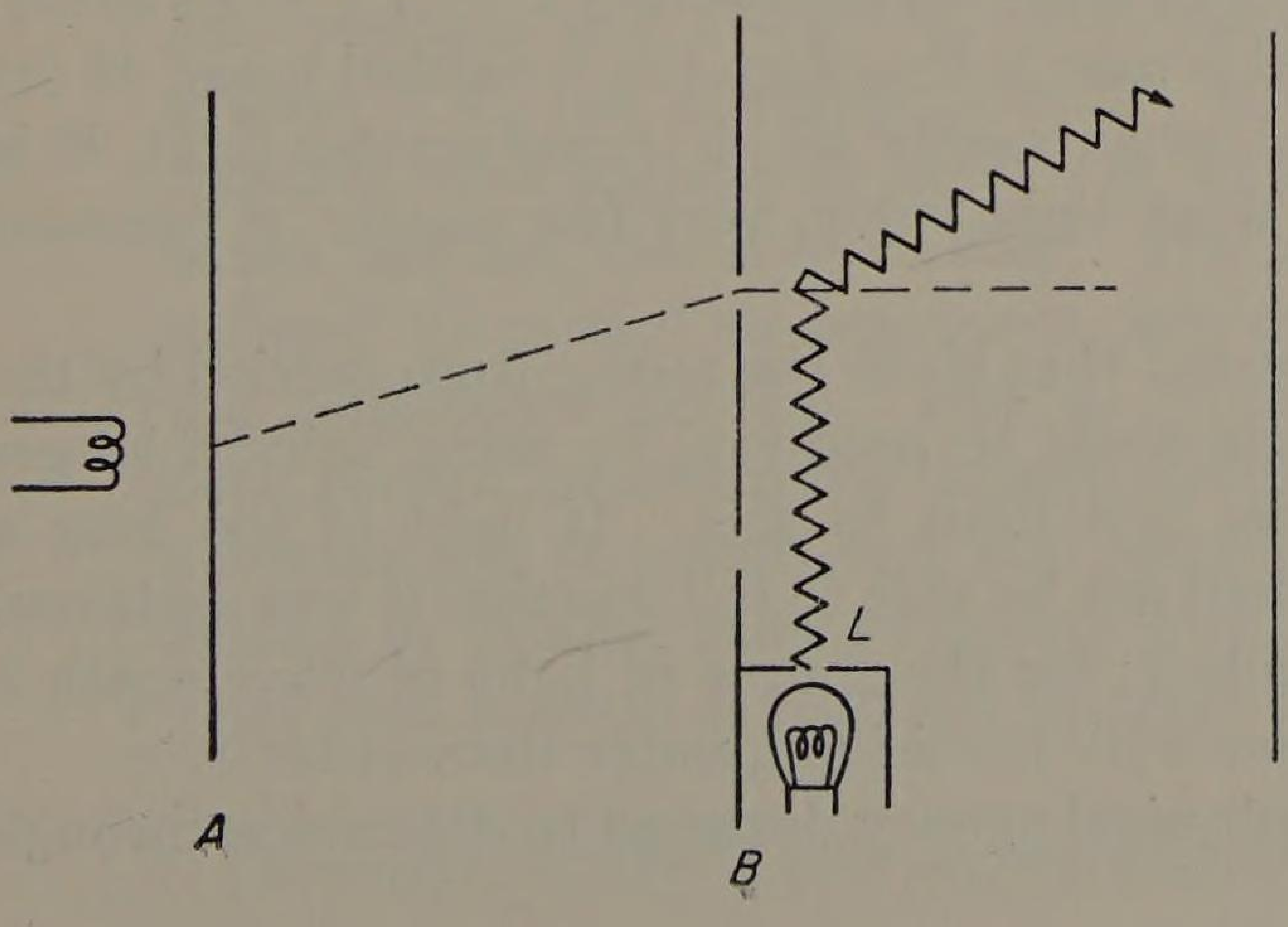


Fig. 1-4 A modification of the experiment of Fig. 1-1. Here we place a light source L behind the screen B and look for light scattered by the electrons passing through hole 1 or hole 2. With a strong light source every electron is indeed found to pass by one or the other hole. But now the probability of arrival at x is no longer given by the curve of Fig. 1-2a, but is instead given by Fig. 1-2d.

indeed, very nearly distributed as in curve (*b*). (This result is obtained whether hole 2 is open or closed, so we have verified that there is no subtle influence of closing 2 on the motion near hole 1.) If we select the electrons scattering light at 2, we get (very nearly) P_2 of Fig. 1-2*c*. But now each electron appears at either 1 or 2 and we can separate our electrons into disjoint classes. So, if we take both together, we *must* get the distribution $P = P_1 + P_2$ illustrated in Fig. 1-2*d*. And experimentally we do! Somehow now the distribution does *not* show the interference effects III of curve (*a*)!

What has been changed? When we watch the electrons to see through which hole they pass, we obtain the result $P = P_1 + P_2$. When we do not watch, we get a different result,

$$P = |\phi_1 + \phi_2|^2 \neq P_1 + P_2$$

Just by watching the electrons, we have changed the chance that they arrive at x. How is this possible? The answer is that, to watch them, we used light and the light in collision with the electron may be expected to alter its motion, or, more exactly, to alter its chance of arrival at x.

On the other hand, can we not use weaker light and thus expect a weaker effect? A negligible disturbance certainly cannot be presumed to produce the finite change in distribution from (*a*) to (*d*). But weak light does not mean a weaker disturbance. Light comes in photons of energy $h\nu$, where ν is the frequency, or of momentum h/λ, where λ is the wavelength. Weakening the light just means using fewer photons, so that we may miss seeing an electron. But when we do see one, it means a complete photon was scattered and a finite momentum of order h/λ is given to the electron.

The electrons that we miss seeing are distributed according to the interference law (*a*), while those we do see and which therefore have scattered a photon arrive at x with the probability $P = P_1 + P_2$ in (*d*). The net distribution in this case is therefore the weighted mean of (*a*) and (*d*). In strong light when nearly all electrons scatter light, it is nearly (*d*); and in very weak light, when very few scatter, it becomes more like (*a*).

It might still be suggested that since the momentum carried by the light is h/λ, weaker effects could be produced by using light of longer wavelength λ. But there is a limit to this. If light of too long a wavelength is used, we shall not be able to tell whether it was scattered from behind hole 1 or hole 2; for the source of light of wavelength λ cannot be located in space with precision greater than order λ.

We thus see that any physical agency designed to determine through which hole the electron passes must produce, lest we have a paradox, enough disturbance to alter the distribution from (*a*) to (*d*).

It was first noticed by Heisenberg, and stated in his uncertainty principle, that the consistency of the then-new mechanics required a limitation to the subtlety to which experiments could be performed. In our case the principle says that an attempt to design apparatus to determine through what hole the electron passed, and delicate enough so as not to deflect the electron sufficiently to destroy the interference pattern, must fail. It is clear that the consistency of quantum mechanics requires that it must be a general statement involving all the agencies of the physical world which might be used to determine through which hole an electron passes. The world cannot be half quantum-mechanical, half classical. No exception to the uncertainty principle has been discovered.

1-2 THE UNCERTAINTY PRINCIPLE

We shall state the uncertainty principle as follows: Any determination of the alternative taken by a process capable of following more than one alternative destroys the interference between alternatives. Heisenberg's original statement of the uncertainty principle was not given in the form we have used here. We shall interrupt our argument for a few paragraphs to discuss Heisenberg's original statement.

In classical physics a particle can be described as moving along a definite trajectory and having, for example, a precise position and velocity at any particular time. Such a picture would not lead to the odd results that we have seen are characteristic of quantum mechanics. Heisenberg's uncertainty principle gives the limits of accuracy of such classical ideas. For example, the idea that a particle has both a definite position and a definite momentum has its limitations. A real system (i.e., one obeying quantum mechanics) looked upon from a classical view appears to be one in which the position or momentum is not definite, but is uncertain. The uncertainty in position can be reduced by careful measurement, and other measurements may make the momentum definite. But, as Heisenberg stated in his principle, both cannot be accurately known simultaneously; the product of the uncertainties of momentum and position involved in any experiment cannot be smaller than a number with the order of magnitude of $\hbar$.†

That such a result is required by physical consistency in the situation we have been discussing can be shown by considering still another way of trying to determine through which hole the electron passes.

---

† $\hbar = h/2\pi = 1.054 \times 10^{-27}$ erg-sec, where h = Planck's constant.

Example. Notice that if an electron is deflected in passing through one of the holes, its vertical component of momentum is changed. Furthermore, an electron arriving at the detector x after passing through hole 1 is deflected by a different amount, and thus suffers a different change in momentum, than an electron arriving at x via hole 2. Suppose that the screen at B is not rigidly supported, but is free to move up and down (Fig. 1-5). Any change in the vertical component of the momentum of an electron upon passing through a hole will be accompanied by an equal and opposite change in the momentum of the screen. This change in momentum can be measured by measuring the velocity of the screen before and after the passage of an electron. Call δp the difference in momentum change between electrons passing through hole 1 or hole 2. Then an unambiguous determination of the hole used by a particular electron requires a momentum determination of the screen to an accuracy of better than δp.

If the experiment is set up in such a way that the momentum of screen B can be measured to the required accuracy, then, since we can

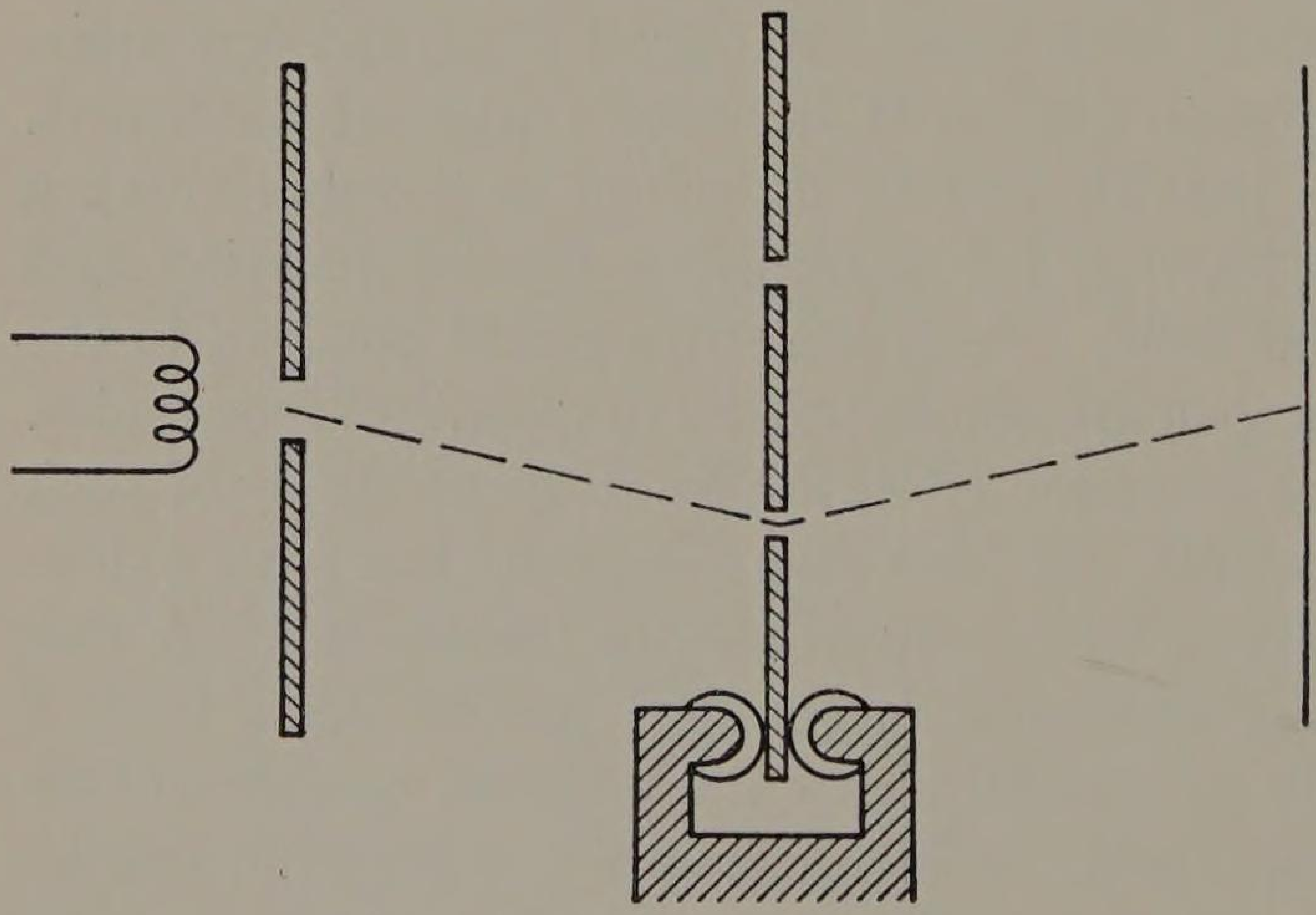


Fig. 1-5 Another modification of the experiment of Fig. 1-1. The screen B is left free to move vertically. If the electron passes hole 2 and arrives at the detector (at $x = 0$, for example), it is deflected upward and the screen B should recoil downward. The hole through which the electron passes can be determined for each passage by starting with the screen at rest and measuring whether it is recoiling up or down afterward. According to Heisenberg's uncertainty principle, however, such precise momentum measurements on screen B would be inconsistent with accurate knowledge of its vertical position, so we could not be sure that the center line of the holes is correctly set. Instead of $P(x)$ of Fig. 1-2a, we get this smeared a little in the vertical direction, so it looks like Fig. 1-2d.

determine the hole passed through, we must find that the resulting distribution of electrons is that of curve (d) of Fig. 1-2. The interference pattern of curve (a) must be lost. How can this happen? To understand, note that the construction of a distribution curve in the plane C requires an accurate knowledge of the vertical position of the two holes in screen B. Thus we must measure not only the momentum of screen B but also its position. If the interference pattern of curve (a) is to be established, the vertical position of B must be known to an accuracy of better than $d/2$, where d is the spacing between maxima of the curve (a). For suppose the vertical position of B is not known to this accuracy; then the vertical position of every point in Fig. 1-2a cannot be specified with an accuracy greater than $d/2$, since the zero point of the vertical scale must be lined up with some nominal zero point on B. Then the value of P at any particular height x must be obtained by averaging over all values within a distance $d/2$ of x. Clearly, the interference pattern will be smeared out by this averaging process. The resulting curve will look like Fig. 1-2d.

The interference pattern in the original experiment is the sign of a wave-like behavior of the electrons. The pattern is the same for any wave motion, so we may use the well-known result from the theory of light diffraction that the relation between the separation a of the holes, the distance l between screen B and the plane C, the wavelength λ of the light, and d is

$$\frac{a}{l} = \frac{\lambda}{d} \tag{1-4}$$

as shown in Fig. 1-6. In Chap. 3 we shall find that the wavelength of the electron waves is intimately connected with the momentum of the electron by the relation

$$p = \frac{h}{\lambda} \tag{1-5}$$

If p is the total momentum of an electron (and we assume all the electrons have the same total momentum), then for $l \gg a$

$$\frac{\delta p}{p} \simeq \frac{a}{l} \tag{1-6}$$

as shown in Fig. 1-7. It follows that

$$d = \frac{h}{\delta p} \tag{1-7}$$

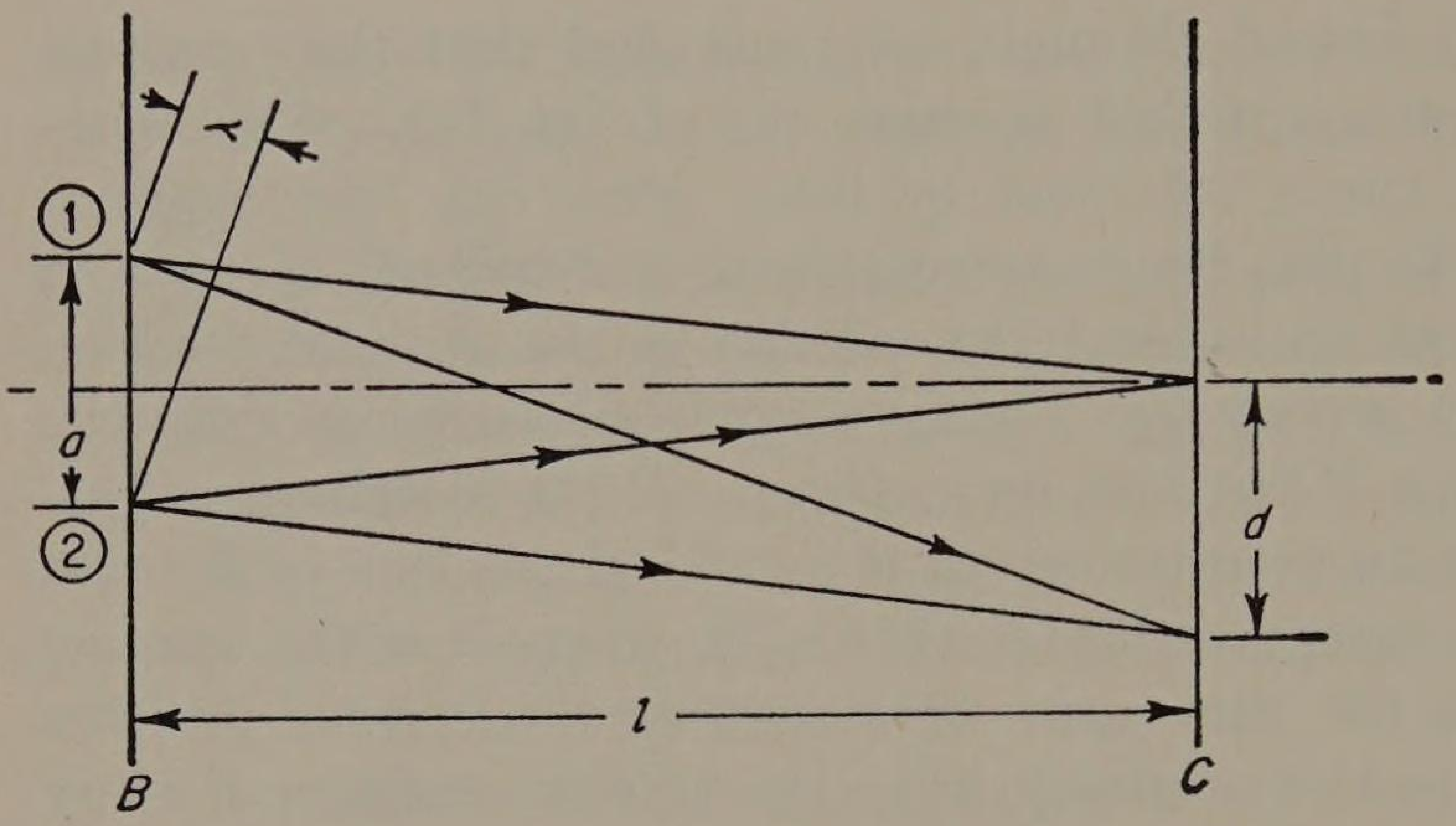


Fig. 1-6 Two beams of light, starting in phase at holes 1 and 2, will interfere constructively when they reach the screen C if they take the same time to travel from B to C. This means that a maximum in the diffraction pattern for light beams passing through two holes will occur at the center of the screen. As we move down the screen, the next maximum will occur at a distance d, which is far enough from the center that, in traveling to this point, the beam from hole 1 will have traveled exactly one wavelength λ farther than the beam from hole 2.

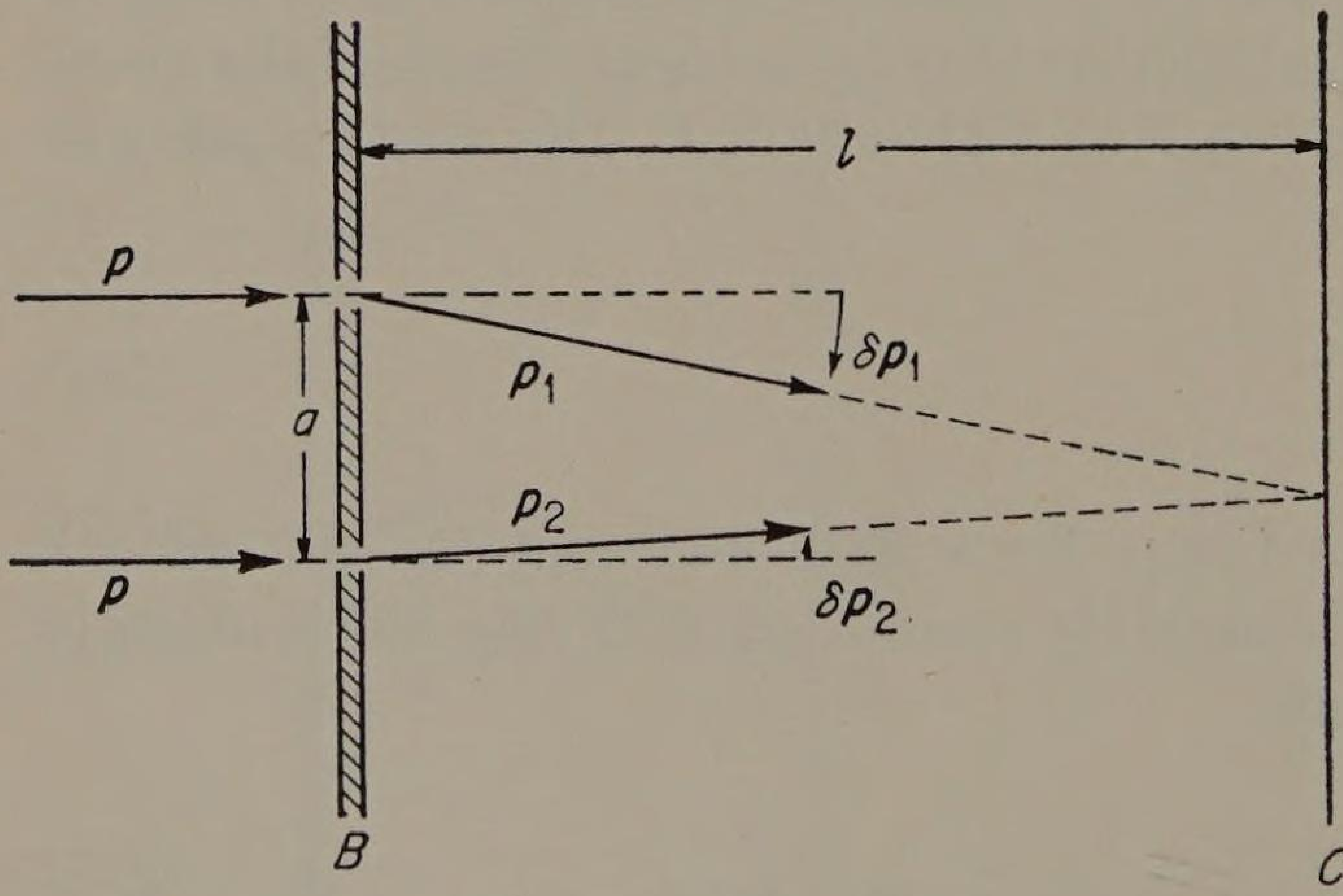


Fig. 1-7 The deflection of an electron in passing through a hole in the screen B is actually a change in momentum $\delta\mathbf{p}$. This change amounts to the addition of a small component of momentum in a direction approximately perpendicular to the original momentum vector. The change in energy is completely negligible. For small deflection angles, the total momentum vector keeps the same magnitude (approximately). Then the deflection angle is represented to a very good approximation by $|\delta\mathbf{p}|/|\mathbf{p}|$. If two electrons, one starting from hole 1 with momentum $\mathbf{p}_1$ and the other starting from hole 2 with momentum $\mathbf{p}_2$, reach the same point on the screen C, then the angles through which they were deflected must differ by approximately a/l. Since we cannot say through which hole an electron has come, the uncertainty in the vertical component of momentum which the electron receives on passing through the screen B must be equivalent to this uncertainty in deflection angle. This gives relation $|\mathbf{p}_1 - \mathbf{p}_2|/|\mathbf{p}| = |\delta\mathbf{p}|/|\mathbf{p}| = a/l$.

Since experimentally we find that the interference pattern has been lost, it must be that the uncertainty δx in the measurement of the position of B is larger than $d/2$. Thus

$$\delta p\ \delta x \geq \frac{h}{2} \tag{1-8}$$

which agrees (in order of magnitude) with the usual statement of the uncertainty principle.

A similar analysis can be applied to the previous measuring device where the scattering of light was used to determine through which hole the electron passed. Such an analysis produces the same lower limit for the uncertainties of measurement.

The uncertainty principle is not "proved" by considering a few such experiments. It is only illustrated. The evidence for it is of two kinds. First, no one has yet found any experimental way to defeat the limitations in measurements which it implies. Second, the laws of quantum mechanics seem to require it if their consistency is to be maintained, and the prediction of these laws has been confirmed again and again with great precision.

1-3 INTERFERING ALTERNATIVES

Two Kinds of Alternatives. From a physical standpoint the two routes are independent alternatives, yet the implication that the probability is the sum $P_1 + P_2$ is false. This means that either the premise or the reasoning which leads to such a conclusion must be false. Since our habits of thought are very strong, many physicists find that it is much more convenient to deny the premise than to deny the reasoning. To avoid the logical inconsistencies into which it is so easy to stumble, they take the following view: When no attempt is made to determine through which hole the electron passes, one cannot say it must pass through one hole or the other. Only in a situation where apparatus is operating to determine through which hole the electron goes is it permissible to say that it passes through one or the other. When you watch, you find that it goes through either one hole or the other hole; but if you are not looking, you cannot say that it goes either one way or the other! Nature demands that we walk a logical tightrope if we wish to describe her.

Contrary to that way of thinking, we shall in this book follow the suggestion made in the first part of this chapter and deny the reasoning; i.e., we shall not compute probabilities by adding probabilities for all alternatives. In order to make definite the new rules for combining probabilities, it will be convenient to define two meanings for the word "alternative." The first of these meanings carries with it the concept of exclusion. Thus holes 1 and 2 are *exclusive alternatives* if one of them is closed or if some apparatus that can unambiguously determine which hole is used is operating. The other meaning of the word "alternative" carries with it a concept of combination or interference. (The term *interference* has the same meaning here as it has in optics, i.e., either constructive or destructive interference.) Thus we shall say, holes 1 and 2 present *interfering alternatives* to the electron when (1) both holes are open and (2) no attempt is made to determine through which hole the electron passes. When the alternatives are of this interfering type, the laws of probability must be changed to the form given in Eqs. (1-1) and (1-2).

The concept of interfering alternatives is fundamental to all of quantum mechanics. In some situations we may have both kinds of alternatives present. Suppose we ask, in the two-hole experiment, for the probability that the electron arrives at some point, say, within 1 cm of the center of the screen. We may mean by that the probability that if there were counters arranged all over the screen (so one or another would go off when the electron arrived), the counter which went off was within 1 cm of $x = 0$. Here the various possibilities are that the electron arrives at some counter via some hole. The holes represent interfering alternatives, but the counters are exclusive alternatives. Thus we first add $\phi_1 + \phi_2$ for a fixed x, square that, and then sum these resultant probabilities over x from -1 to $+1$ cm.

It is not hard, with a little experience, to tell which kind of alternatives is involved. For example, suppose that information about the alternatives is available (or could be made available without altering the result), but this information is not used. Nevertheless, in this case a sum of probabilities (in the ordinary sense) must be carried out over *exclusive* alternatives. These exclusive alternatives are those which *could* have been separately identified by the information.

Some Illustrations. When alternatives cannot possibly be resolved by any experiment, they always interfere. A striking illustration of this is the scattering of two nuclei at 90°, say, in the center-of-gravity system, as illustrated in Fig. 1-8. Suppose A represents an α particle and B some other nucleus. Ask for the probability that A is scattered to position 1 and B to 2. The amplitude is, say, $\phi_{AB}(1,2)$. The probability of this is $p = |\phi_{AB}(1,2)|^2$. Suppose we do not distinguish what kind of nucleus arrives at 1, that is, whether it

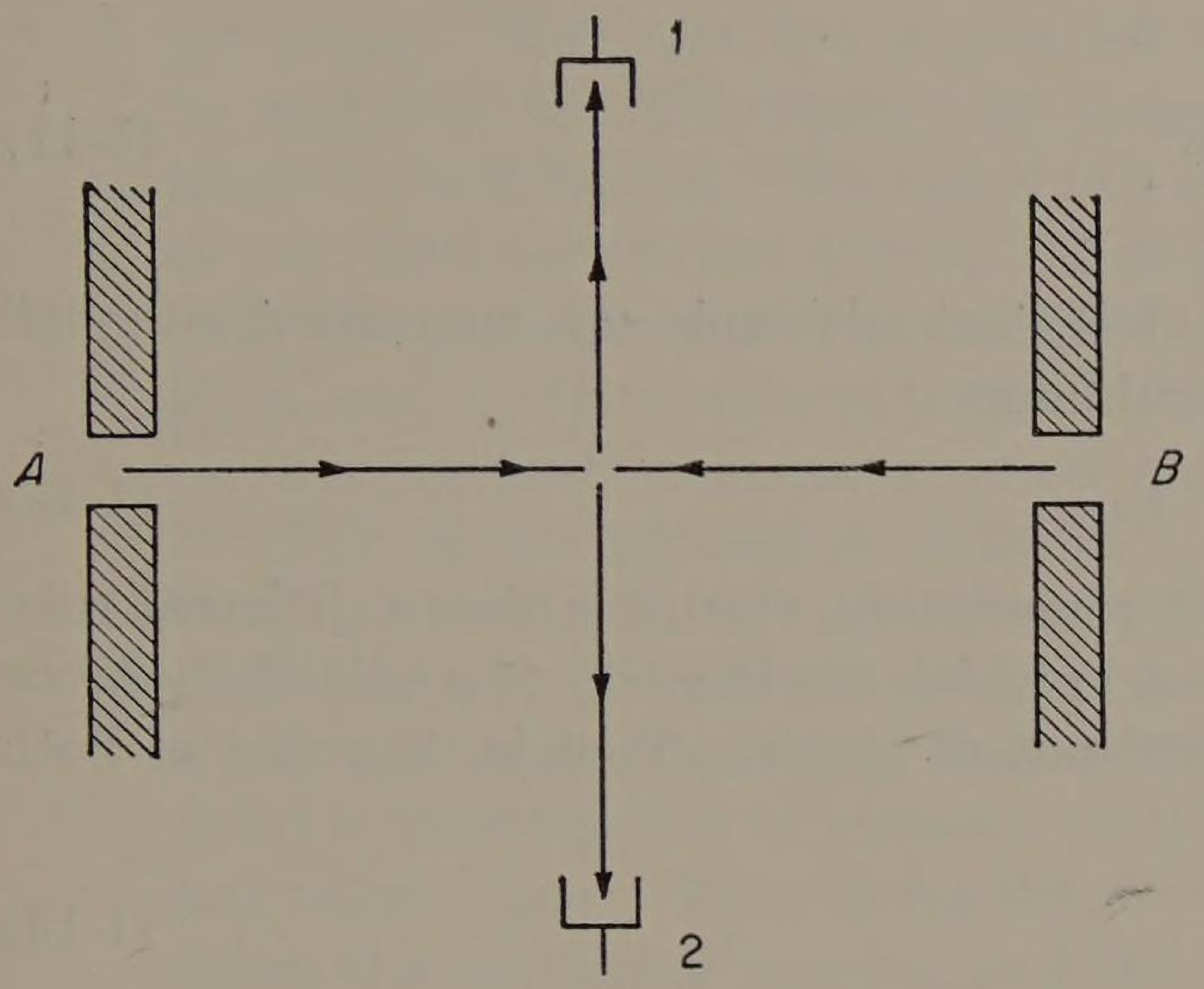


Fig. 1-8 Scattering of one nucleus by another in the center-of-gravity system. The scattering of two identical nuclei shows striking interference effects. There are two interfering alternatives here. The particle which arrives at 1, say, can have been that which started either from A or from B. If the original nuclei were not identical, tests of identity at 1 could determine which alternative had actually been taken, so they are exclusive alternatives and the special interference effects do not arise in this case.

is B or A. If it is B, the amplitude is $\phi_{AB}(2,1)$ [which equals $\phi_{AB}(1,2)$, because we have taken a 90° angle]. The chance that some nucleus arrives at 1 and the other at 2 is

$$|\phi_{AB}(1,2)|^2 + |\phi_{AB}(2,1)|^2 = 2p \tag{1-9}$$

We have added the probabilities. The cases A arrives at 1 and B arrives at 1 are exclusive alternatives because we could, if we wished, determine the character of the nucleus at 1 without disturbing the previous scattering process.

But if A is an α particle, what happens if B is also an α particle? Then no experiment can distinguish which is which, and we cannot know when one arrives at 1 whether it is A or B. We have interfering alternatives, and the probability is

$$|\phi_{AB}(1,2) + \phi_{AB}(2,1)|^2 = 4p \tag{1-10}$$

This interesting result is readily verified experimentally.

If electrons scatter electrons, the result is different in two ways. First, the electron has a quality we call *spin*, and a given electron may be in one of two states called *spin up* and *spin down*. The spin is not changed to first approximation for scattering at low energy. The spin carries a magnetic moment. At low velocities the main forces are electrical, owing to charge, and the magnetic influences make only a small correction, which we neglect. So if electron A has spin up and B has spin down, we could later tell which arrived at 1 by measuring its spin. If up, it is A; if down, it is B. The scattering probability is then

$$|\phi(1,2)|^2 + |\phi(2,1)|^2 = 2p \tag{1-11}$$

in this case.

If, however, both A and B start with spin up, we cannot later tell which is which and we would expect

$$|\phi_{AB}(1,2) + \phi_{AB}(2,1)|^2 = 4p \tag{1-12}$$

Actually this is wrong and, remarkably, electrons obey a different rule. The amplitude for an event in which the identity of a pair of electrons is reversed contributes 180° out of phase. That is, the case of both spin up gives

$$|\phi_{AB}(1,2) - \phi_{AB}(2,1)|^2 \tag{1-13}$$

In our case of 90° scattering $\phi_{AB}(1,2) = \phi_{AB}(2,1)$, so this is zero.

Fermions and Bosons. This rule of the 180° phase shift for alternatives involving exchange in identity of electrons is very odd, and its ultimate reason in nature is still only imperfectly understood. Other particles besides electrons obey it. Such particles are called fermions, and are said to obey Fermi, or antisymmetric, statistics. Electrons, protons, neutrons, neutrinos, and μ mesons are fermions. So are compounds of an odd number of these such as a nitrogen atom, which contains seven electrons, seven protons, and seven neutrons. This 180° rule was first stated by Pauli and is the full quantum-mechanical basis of his exclusion principle, which controls the character of the chemists' periodic table.

Particles for which interchange does not alter the phase are called bosons and are said to obey Bose, or symmetrical, statistics. Examples of bosons are photons, π mesons, and systems containing an even number of Fermi particles such as an α particle, which is two protons and two neutrons. All particles are either one or the other, bosons or fermions. These interference properties can have profound and mysterious effects. For example, helium liquid made of atoms of atomic mass 4 (bosons) at temperatures of one or two degrees Kelvin can flow without any resistance through small tubes, whereas the liquid made of atoms of mass 3 (fermions) does not have this property.

The concept of identity of particles is far more complete and definite in quantum mechanics than it is in classical mechanics. Classically, two particles which seem identical could be nearly identical, or identical for all practical purposes, in the sense that they may be so closely equal that present experimental techniques cannot detect any difference. However, the door is left open for some future technique to establish the difference. In quantum mechanics, however, the situation is different. We can give a direct test to determine whether or not particles are completely indistinguishable.

If the particles in the experiment diagramed in Fig. 1-8, starting at A and B, were only approximately identical, then improvements in experimental techniques would enable us to determine by close scrutiny of the particle arriving at 1, for example, whether it came from A or B. In this situation the alternatives of the two initial positions must be exclusive, and there must be no interference between the amplitudes describing these alternatives. Now the important point is that this act of scrutiny would take place *after* the scattering had taken place. This means that the observation could not possibly affect the scattering process, and this in turn implies that we would expect no interference between the amplitudes describing the alternatives (that it is either the particle from A or the particle from B which arrives at 1). In this case we must conclude from the uncertainty principle that there is no way, even in principle, to ever distinguish between these possibilities. That is, when a particle arrives at 1, it is completely impossible by any test whatsoever, now or in the future, to determine whether the particle started from A or B. In this more rigorous sense of identity, all electrons are identical, as are all protons, etc.

As a second example we consider the scattering of neutrons from a crystal. When neutrons of wavelength somewhat shorter than the atomic spacing are scattered from the atoms in a crystal, we get very strong interference effects. The neutrons emerge only in certain discrete directions determined by the Bragg law of reflection, just as for X rays. The interfering alternatives which enter this example are the alternative possibilities that it is this, or that, atom which does the scattering of a particular neutron. (The amplitude to scatter neutrons from any atom is so small that we need not consider alternatives in which a neutron is scattered by more than one atom.) The waves of amplitude describing the motion of a neutron which start from these atoms interfere constructively only in certain definite directions.

Now there is an interesting complication which enters this apparently simple picture. Neutrons, like electrons, carry a spin, which can be analyzed in two states, spin up and spin down. Suppose the scattering material is composed of an atomic species which has a similar spin property, such as carbon-13. In this case an experiment will reveal two apparently different types of scattering. It is found that besides the scattering in discrete directions, as described in the preceding paragraph, there is a diffused scattering in all directions. Why should this be?

A clue to the source of these two types of scattering is provided by the following observation. Suppose all the neutrons which enter the experiment are prepared in such a manner that their spin direction is up. If the spin direction of the emerging neutrons is analyzed, it will be found that there are some up and some down; those which still have spin up are scattered only at the discrete Bragg angles, while those whose spin has been changed to down come out scattered diffusely in all directions!

Now in order that a neutron flip its spin from up to down, the law of conservation of angular momentum requires that the spin of the scattering nucleus be changed from down to up. Therefore, in principle, the particular nucleus which was responsible for scattering that particular neutron could be determined. We could, in principle, note down before the experiment the spin state of all the scattering nuclei in the crystal. Then, after the neutron is scattered, we could reinvestigate the crystal and see which nucleus had changed its spin from down to up. If no crystal nucleus underwent such a change in spin, then neither did the neutron, and we cannot tell from which nucleus the neutron was actually scattered. In this case the alternatives interfere and the Bragg law of scattering results.

If, on the other hand, one crystal nucleus is found to have changed spin, then we know that this nucleus did the scattering. There are no interfering alternatives. The spherical waves of amplitude which emerge from this particular nucleus describe the motion of the scattered neutron, and only the waves emerging from this nucleus enter into that description. In this case there is equal likelihood to find the scattered neutron coming out in any direction.

The concept of searching through all the nuclei in a crystal to find which one has changed its spin state is surely a needle-in-the-haystack type of activity, but nature is not concerned with the practical difficulties of experimentation. The important fact is that in principle it is possible without producing any disturbance of the scattered neutron to determine (in this latter case where the spin states change) which crystal nucleus actually did the scattering. The existence of this possibility means that even if we do not actually carry out this determination, we are nevertheless dealing with exclusive (and thus noninterfering) alternatives.

On the other hand, the fact that we get interference between alternatives in the situation where the spin states of the neutrons were not changed means that it is impossible, even in principle, to ever discover which particular crystal nucleus did the scattering—impossible, at least, without disturbing the situation during or before the scattering.

1-4 SUMMARY OF PROBABILITY CONCEPTS

Alternatives and the Uncertainty Principle. The purpose of this introductory chapter has been to explain the meaning of a probability amplitude and its importance in quantum mechanics and to discuss the rules for manipulation of these amplitudes. Thus we have stated that there is a quantity called a *probability amplitude* associated with every method whereby an event in nature can take place. For example, an electron going from a source at S (Fig. 1-1) to a detector at x has one amplitude for completing this course while passing through hole 1 of the screen at B and another amplitude in passing through hole 2. Further, we can associate an amplitude with the overall event by adding together the amplitudes of each alternative method. Thus, for example, the overall amplitude for arrival at x is given in Eq. (1-2) as

$$\phi = \phi_1 + \phi_2 \tag{1-14}$$

Next, we interpret the absolute square of the overall amplitude as the probability that the event will happen. For example, the probability that an electron reaches the detector is

$$P = |\phi_1 + \phi_2|^2 \tag{1-15}$$

If we interrupt the course of the event before its conclusion with an observation on the state of the particles involved in the event, we disturb the construction of the overall amplitude. Thus if we observe the system of particles to be in one particular state, we exclude the possibility that it can be in any other state, and the amplitudes associated with the excluded states can no longer be added in as alternatives in computing the overall amplitude. For example, if we determine with the help of some sort of measuring equipment that the electron passes through hole 1, the amplitude for arrival at the detector is just ϕ_1. Further, it does not matter if we actually observe and record the outcome of the measurement or not, so long as the measuring equipment is working. Obviously, we could observe the outcome at any time we wished. The operation of the measuring equipment is sufficient to disturb the system and its probability amplitude.

This latter fact is the basis of the Heisenberg uncertainty principle, which states that there is a natural limit to the subtlety of any experiment or the refinement of any measurement.

The Structure of the Amplitude. The amplitude for an event is the sum of the amplitudes for the various alternative ways that the event can occur. This permits the amplitude to be analyzed in many different ways depending on the different classes into which the alternatives can be divided. The most detailed analysis results from considering that a particle going from A to B, for example, in a given time interval, can be considered to have done this by going in a certain motion (position vs. time) or path in space and time. We shall therefore associate an amplitude with each possible motion. The total amplitude will be the sum of a contribution from each of the paths.

This idea can be made more clear by a further consideration of our experiment with the two holes. Suppose we put a couple of extra screens between the source and the hole. Call these screens D and E. In each of them we drill a few holes which we number D_1, D_2, . . . and E_1, E_2, . . . (Fig. 1-9). For simplicity, we shall assume the electrons are constrained to move in the xy plane. Then there are several alternative paths which an electron may take in going from the source to the hole in screen B. It could go from source to D_2, and then E_3, and then the hole 1; or it could go from the source to D_3, then E_1, and finally to the hole 1, etc. Each of these paths has its own amplitude. The complete amplitude is the sum of all of them.

Next, suppose we continue to drill holes in the screens D and E until there is nothing left of the screens. The path of an electron must now be specified by the height x_D at which the electron passes

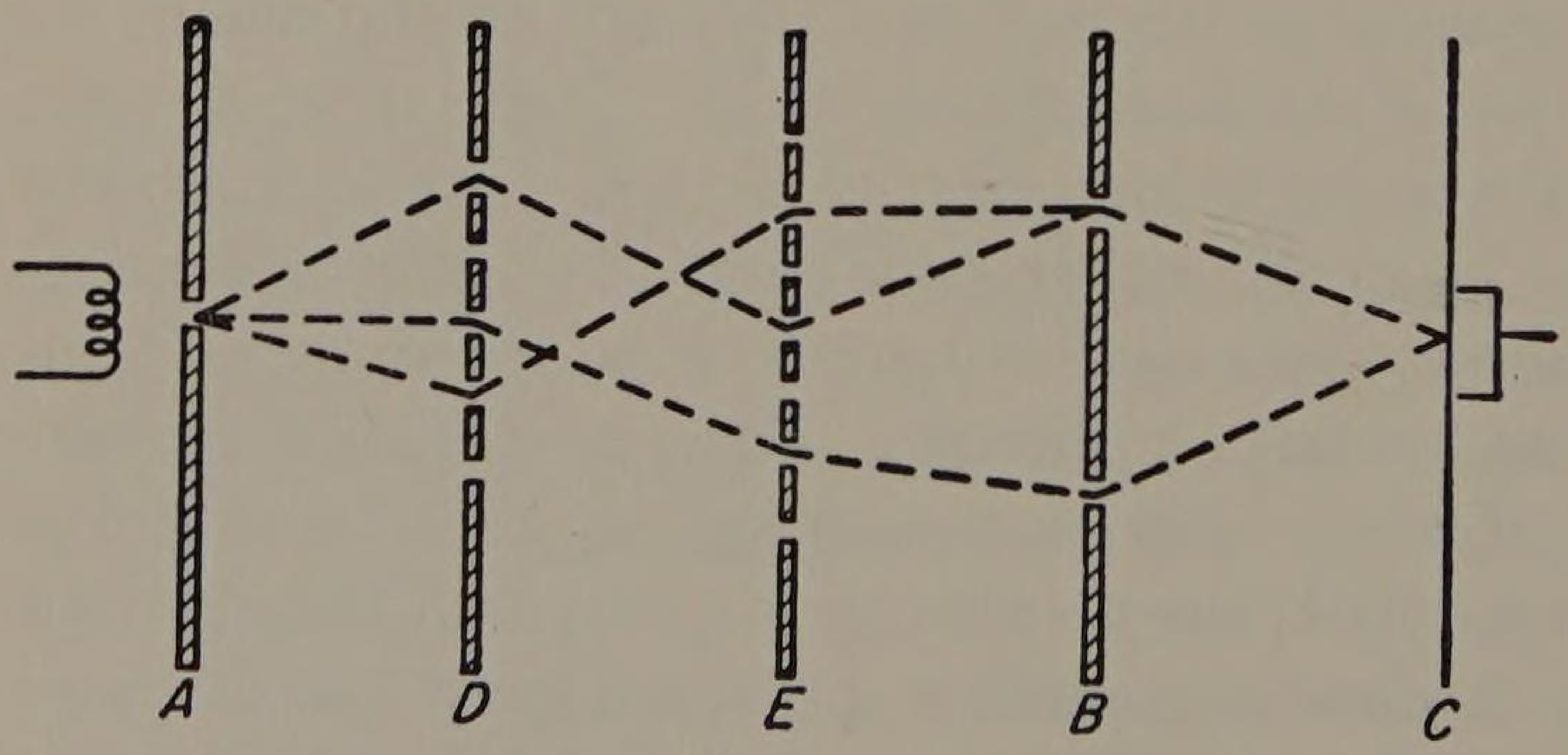


Fig. 1-9 When several holes are drilled in the screens D and E placed between the source in screen A and the final position in screen C, several alternative routes are available for each electron. For each of these routes there is an amplitude. The result of any experiment in which all of the holes are open requires the addition of all of these amplitudes, one for each possible path.

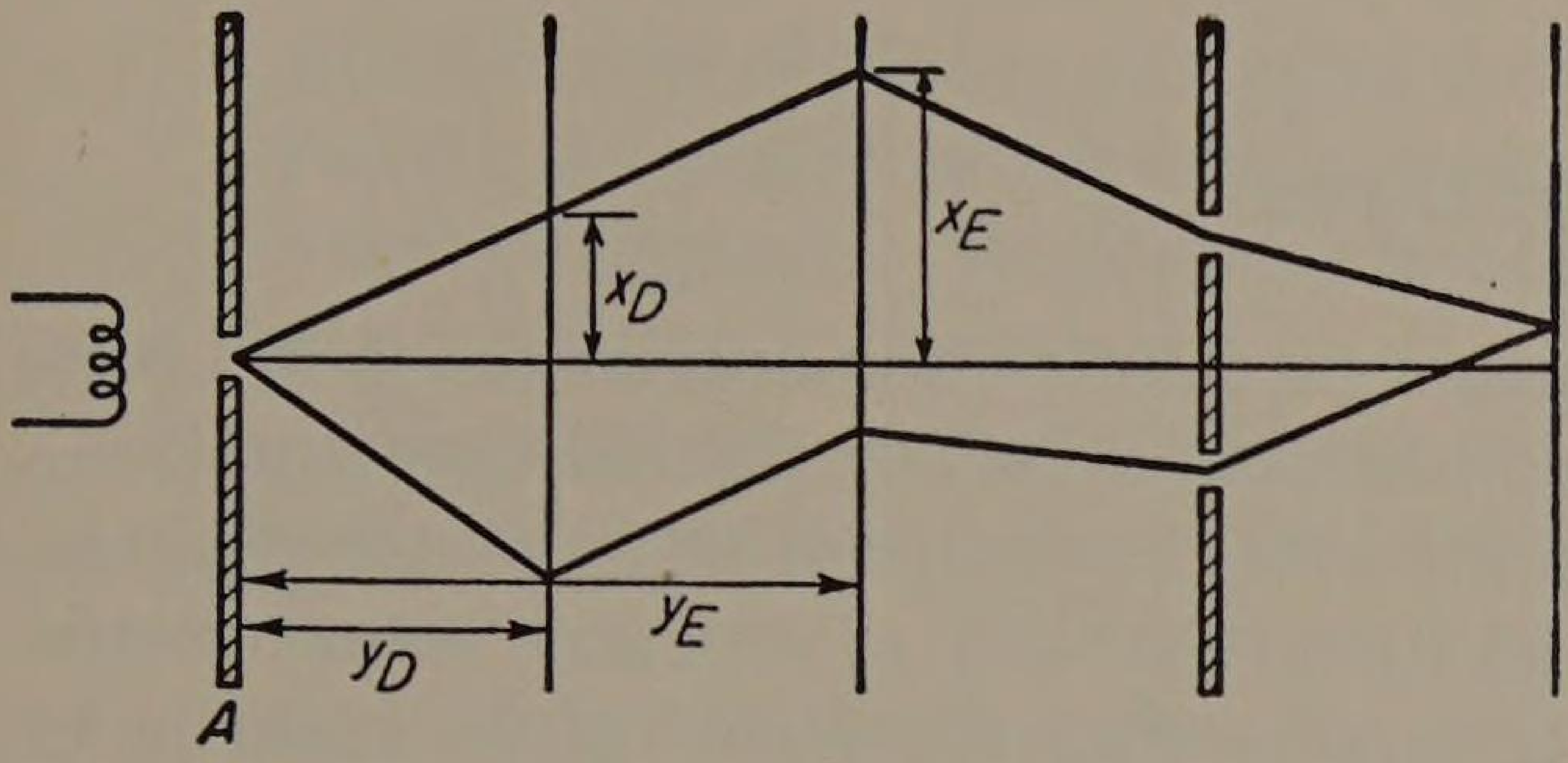


Fig. 1-10 More and more holes are cut in the screens at y_D and y_E. Eventually, the screens are completely riddled with holes, and the electron has a continuous range of positions, up and down along each screen, at which it can pass through the position of the screen. In this case the sum of alternatives becomes a double integral over the continuous parameters x_D and x_E describing the alternative heights at which the electron passes the position of the screens at y_D and y_E.

the position y_D at the nonexistent screen D, together with the height x_E, at the position y_E, as in Fig. 1-10. To each pair of heights there corresponds an amplitude. The principle of superposition still applies, and we must take the sum (or by now, the integral) of these amplitudes over all possible values of x_D and x_E.

Clearly, the next thing to do is to place more and more screens between the source and the hole 1 and in each screen drill so many holes that there is nothing left. Throughout this process we continue to refine the definition of the path of the electron, until finally we arrive at the sensible idea that a path is merely height as a particular function of distance, or $x(y)$. We also continue to apply the principle of superposition, until we arrive at the integral over all paths of the amplitude for each path.

Now we can make a still finer specification of the motion. Not only can we think of the particular path $x(y)$ in space, but we can specify the *time* at which it passes each point in space. That is, a path will (in our two-dimensional case) be given if the two functions $x(t)$, $y(t)$ are given. Thus we have the idea of an amplitude to take a certain path $x(t)$, $y(t)$. The total amplitude to arrive is the sum or integral of this amplitude over all possible paths. The problem of defining this concept of a sum or integral over all paths in a mathematically more precise way will be taken up in Chap. 2.

Chapter 2 also contains the formula for the amplitude for any given path. Once this is given, the laws of nonrelativistic quantum mechanics are completely stated, and all that remains is a demonstration of the application of these laws in a number of interesting special cases.

1-5 SOME REMAINING THOUGHTS

We shall find that in quantum mechanics, the amplitudes ϕ are solutions of a completely deterministic equation (the Schrödinger equation). Knowledge of ϕ at $t = 0$ implies its knowledge at all subsequent times. The interpretation of $|\phi|^2$ as the probability of an event is an indeterministic interpretation. It implies that the result of an experiment is not exactly predictable. It is very remarkable that this interpretation does not lead to any inconsistencies. That it is true has been amply demonstrated by analyses of many particular situations by Heisenberg, Bohr, Born, von Neumann, and many other physicists. In spite of all these analyses the fact that no inconsistency can arise is not thoroughly obvious. For this reason quantum mechanics appears as a difficult and somewhat mysterious subject to a beginner. The mystery gradually decreases as more examples are tried out, but one never quite loses the feeling that there is something peculiar about the subject.

There are a few interpretational problems on which work may still be done. They are very difficult to state until they are completely worked out. One is to show that the probability interpretation of ϕ is the *only* consistent interpretation of this quantity. We and our measuring instruments are part of nature and so are, in principle, described by an amplitude function satisfying a deterministic equation. Why can we only predict the probability that a given experiment will lead to a definite result? From what does the uncertainty arise? Almost without doubt it arises from the need to amplify the effects of single atomic events to such a level that they may be readily observed by large systems. The details of this have been analyzed only on the assumption that $|\phi|^2$ is a probability, and the consistency of this assumption has been shown. It would be an interesting problem to show that *no other* consistent interpretation can be made.

Other problems which may be further analyzed are those dealing with the theory of knowledge. For example, there seems to be a lack of symmetry in time in our knowledge. Our knowledge of the past is qualitatively different from that of the future. In what way is only the probability of a future event accessible to us, whereas the certainty of a past event can often apparently be asserted? These matters again have been analyzed to a great extent. Possibly a little more can be said to clarify the situation, however. Obviously, we are again involved in the consequences of the large size of ourselves and of our measuring equipment. The usual separation of observer and observed which is now needed in analyzing measurements in quantum mechanics should not really be necessary, or at least should be even more thoroughly analyzed. What seems to be needed is the statistical mechanics of amplifying apparatus.

The analyses of such problems are, of course, in the nature of philosophical questions. They are not necessary for the further development of physics. We know we have a consistent interpretation of ϕ and, almost without doubt, the only consistent one. The problem of today seems to be the discovery of the laws governing the behavior of ϕ for phenomena involving nuclei and mesons. The interpretation of ϕ is interesting. But the much more intriguing question is: What new modifications of our thinking will be required to permit us to analyze phenomena occurring within nuclear dimensions?

1-6 THE PURPOSE OF THIS BOOK

So far, we have given the form the quantum-mechanical laws must take, i.e., that a probability amplitude exists, and we have outlined one possible scheme for calculating this amplitude. There are other ways to formulate this. In a more usual approach to quantum mechanics the amplitude is calculated by solving a kind of wave equation. For particles of low velocity, it is called the Schrödinger equation. A more accurate equation valid for electrons of velocity arbitrarily close to the velocity of light is the Dirac equation. In this case the probability amplitude is a kind of hypercomplex number. We shall not discuss the Dirac equation in this book, nor shall we investigate the effects of spin. Instead, we limit our attention to low-velocity electrons, extending our horizon somewhat in the direction of quantum electrodynamics by investigating photons, particles whose behavior is determined by Maxwell's equation.

In this book we shall give the laws to compute the probability amplitude for nonrelativistic problems in a manner which is somewhat unconventional. In some ways, particularly in developing a conceptual understanding of quantum mechanics, it may be preferred, but in others, e.g., in making computations for the simpler problems and for understanding the literature, it is disadvantageous.

The more conventional view, via the Schrödinger equation, is already presented in many books, but the views to be presented here have appeared only in abbreviated form in papers in the journals.[1]

---

[1] R. P. Feynman, The Space-Time Approach to Non-relativistic Quantum Mechanics, *Rev. Mod. Phys.*, vol. 20, p. 367, 1948.

A central aim of this book is to collect this work into one volume where it may be expanded with sufficient clarity and detail to be of use to the interested student.

In order to keep the subject within bounds, we shall not make a complete development of quantum mechanics. Instead, whenever a topic has reached such a point that further elucidation would best be made by conventional arguments appearing in other books, we refer to those books. Because of this incompleteness, this book cannot serve as a complete textbook of quantum mechanics. It can serve as an introduction to the ideas of the subject if used in conjunction with another book that deals with the Schrödinger equation, matrix mechanics, and applications of quantum mechanics.

On the other hand, we shall use the space saved (by our not developing all of quantum mechanics in detail) to consider the application of the mathematical methods used in the formulation of quantum mechanics to other branches of physics.

It is a problem of the future to discover the exact manner of computing amplitudes for processes involving the apparently more complicated particles, namely, neutrons, protons, and mesons. Of course, one can doubt that, when the unknown laws are discovered, we shall find ourselves computing amplitudes at all. However, the situation today does not seem analogous to that preceding the discovery of quantum mechanics.

In the 1920's there were many indications that the fundamental theorems and concepts of classical mechanics were wrong, i.e., there were many paradoxes. General laws could be proved independently of the detailed forces involved. Some of these laws did not hold. For example, each spectral line showed a degree of freedom for an atom, and at temperature T each degree of freedom should have an energy kT, contributing R to the specific heat. Yet this very high specific heat expected from the enormous number of spectral lines did not appear.

Today, any general law that we have been able to deduce from the principle of superposition of amplitudes, such as the characteristics of angular momentum, seems to work. But the detailed interactions still elude us. This suggests that amplitudes will exist in a future theory, but their method of calculation may be strange to us.

2

The Quantum-mechanical Law of Motion

In this chapter we intend to complete our specification of nonrelativistic quantum mechanics which we began in Chap. 1. There we noted the existence of an amplitude for each trajectory; here we shall give the form of the amplitude for each trajectory. For a while, for simplicity, we shall restrict ourselves to the case of a particle moving in one dimension. Thus the position at any time can be specified by a coordinate x, a function of t. By the path, then, we mean a function $x(t)$.

If a particle at an initial time t_a starts from the point x_a and goes to a final point x_b at time t_b, we shall say simply that the particle goes from a to b and our function $x(t)$ will have the property that $x(t_a) = x_a$ and $x(t_b) = x_b$. In quantum mechanics, then, we shall have an amplitude, often called a *kernel*, which we may write $K(b,a)$, to get from the point a to the point b. This will be the sum over all of the trajectories that go between the end points a and b of a contribution from each. This is to be contrasted with the situation in classical mechanics in which there is only one specific and particular trajectory which goes from a to b, the so-called *classical trajectory*, which we shall label $\bar{x}(t)$. Before we go on to give the rule for the quantum-mechanical case, let us remind ourselves of the situation in classical mechanics.

2-1 THE CLASSICAL ACTION

One of the most elegant ways of expressing the condition that determines the particular path $\bar{x}(t)$ out of all the possible paths is the *principle of least action*. That is, there exists a certain quantity S which can be computed for each path. The classical path $\bar{x}$ is that for which S is a minimum. Actually, the real condition is that S be merely an extremum. That is to say, the value of S is unchanged in the first order if the path $\bar{x}(t)$ is modified slightly.

The quantity S is given by the expression

$$S = \int_{t_a}^{t_b} L(\dot{x},x,t)\,dt \tag{2-1}$$

where L is the lagrangian for the system. For a particle of mass m moving in a potential $V(x,t)$, which is a function of position and time, the lagrangian is

$$L = \frac{m}{2}\dot{x}^2 - V(x,t) \tag{2-2}$$

The form of the extremum path $\bar{x}(t)$ is determined with the usual procedures of the calculus of variations. Thus, suppose the path is varied away from $\bar{x}$ by an amount $\delta x(t)$; the condition that the end points of $\bar{x}$ are fixed requires

$$\delta x(t_a) = \delta x(t_b) = 0 \tag{2-3}$$

The condition that $\bar{x}$ be an extremum of S means

$$\delta S = S[\bar{x} + \delta x] - S[\bar{x}] = 0 \tag{2-4}$$

to first order in δx. Using the definition of Eq. (2-1) we may write

$$\begin{aligned} S[x + \delta x] &= \int_{t_a}^{t_b} L(\dot{x} + \delta\dot{x}, x + \delta x, t)\, dt \\ &= \int_{t_a}^{t_b} \left[L(\dot{x},x,t) + \delta\dot{x}\, \frac{\partial L}{\partial \dot{x}} + \delta x\, \frac{\partial L}{\partial x} \right] dt \\ &= S[x] + \int_{t_a}^{t_b} \left(\delta\dot{x}\, \frac{\partial L}{\partial \dot{x}} + \delta x\, \frac{\partial L}{\partial x} \right) dt \end{aligned} \tag{2-5}$$

Upon integration by parts, the variation in S becomes

$$\delta S = \delta x\, \frac{\partial L}{\partial \dot{x}} \bigg|_{t_a}^{t_b} - \int_{t_a}^{t_b} \delta x \left[\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} \right] dt \tag{2-6}$$

Since δx is 0 at the end points, the first term on the right-hand side of this equation is 0. Between the end points δx can take on any arbitrary value. Thus the extremum is that curve along which the following condition is always satisfied:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \tag{2-7}$$

This is, of course, the classical lagrangian equation of motion.

In classical mechanics, the *form* of the action integral $S = \int L\, dt$ is interesting, not just the extreme value S_{cl}. This interest derives from the necessity to know the action along a set of neighboring paths in order to determine the path of least action.

In quantum mechanics both the form of the integral and the value of the extremum are important. In the following problems we shall evaluate the extremum in a variety of situations.

Problem 2-1 For a free particle $L = m\dot{x}^2/2$. Show that the action S_{cl} corresponding to the classical motion of a free particle is

$$S_{cl} = \frac{m}{2} \frac{(x_b - x_a)^2}{t_b - t_a} \tag{2-8}$$

Problem 2-2 For a harmonic oscillator $L = (m/2)(\dot{x}^2 - \omega^2 x^2$. With T equal to $t_b - t_a$ show that the classical action is

$$S_{cl} = \frac{m\omega}{2 \sin \omega T} (x_a^2 + x_b^2) \cos \omega T - 2x_a x_b \tag{2-9}$$

Problem 2-3 Find S_{cl} for a particle under a constant force F, that is, $L = m\dot{x}^2/2 - Fx$.

Problem 2-4 Classically, the momentum is defined as

$$p = \frac{\partial L}{\partial \dot{x}} \tag{2-10}$$

Show that the momentum at an end point is

$$\left(\frac{\partial L}{\partial \dot{x}}\right)_{x=x_a} = \frac{\partial S_{cl}}{\partial x_a} \tag{2-11}$$

Hint: Consider the effect on Eq. (2-6) of a change in end points.

Problem 2-5 Classically, the energy is defined as

$$E = L - \dot{x}p \tag{2-12}$$

Show that the energy at an end point is

$$L(x_b) - \dot{x}_b \left(\frac{\partial L}{\partial \dot{x}}\right)_{x=x_b} = \frac{\partial S_{cl}}{\partial t_b} \tag{2-13}$$

Hint: A change in the time of an end point requires a change in path, since all paths must be classical paths.

2-2 THE QUANTUM - MECHANICAL AMPLITUDE

Now we can give the quantum-mechanical rule. We must say how much each trajectory contributes to the total amplitude to go from a to b. It is not that just the particular path of extreme action contributes; rather, it is that all the paths contribute. They contribute equal amounts to the total amplitude, but contribute at different phases. The phase of the contribution from a given path is the action S for that path in units of the quantum of action $\hbar$. That is, to summarize: The probability $P(b,a)$ to go from a point x_a at the time t_a to the point x_b at t_b is the absolute square $P(b,a) = |K(b,a)|^2$ of an amplitude $K(b,a)$ to go from a to b. This amplitude is the sum of contribution $\phi[x(t)]$ from each path.

$$K(b,a) = \sum_{\substack{\text{over all paths}\\ \text{from } a \text{ to } b}} \phi[x(t)] \tag{2-14}$$

The contribution of a path has a phase proportional to the action S:

$$\phi[x(t)] = \text{const } e^{(i/\hbar)S[x(t)]} \tag{2-15}$$

The action is that for the corresponding classical system [see Eq. (2-1)]. The constant will be chosen to normalize K conveniently, and it will be taken up later when we discuss more mathematically just what we mean in Eq. (2-14) by a sum over paths.

2-3 THE CLASSICAL LIMIT

Before we go on to making the mathematics more complete, we shall compare this quantum law with the classical rule. At first sight, from Eq. (2-15) all paths contribute equally, although their phases vary, so it is not clear how, in the classical limit, some particular path becomes most important. The classical approximation, however, corresponds to the case that the dimensions, masses, times, etc., are so large that S is enormous in relation to $\hbar$ ($= 1.05 \times 10^{-27}$ erg-sec). Then the phase of the contribution $S/\hbar$ is some very, very large angle. The real (or imaginary) part of ϕ is the cosine (or sine) of this angle. This is as likely to be plus as minus. Now if we move the path as shown in Fig. 2-1 by a small amount δx, *small on the classical scale*, the change in S is likewise small on the classical scale, but not when measured in the tiny unit $\hbar$. These small changes in path will, generally, make enormous changes in phase, and our cosine or sine will oscillate exceedingly rapidly between plus and minus values. The total contribution will then add to zero; for if one path makes a positive contribution, another infinitesimally close (on a classical scale) makes an equal negative contribution, so that no net contribution arises.

Therefore, no path really needs to be considered if the neighboring path has a different action; for the paths in the neighborhood cancel out the contribution. But for the special path $\bar{x}$, for which S is an extremum, a small change in path produces, in the first order at least, no change in S. All the contributions from the paths in this region are nearly in phase, at phase S_{cl}, and do not cancel out. Therefore, only for paths in the vicinity of $\bar{x}$ can we get important contributions, and in the classical limit we need only consider this particular trajec-

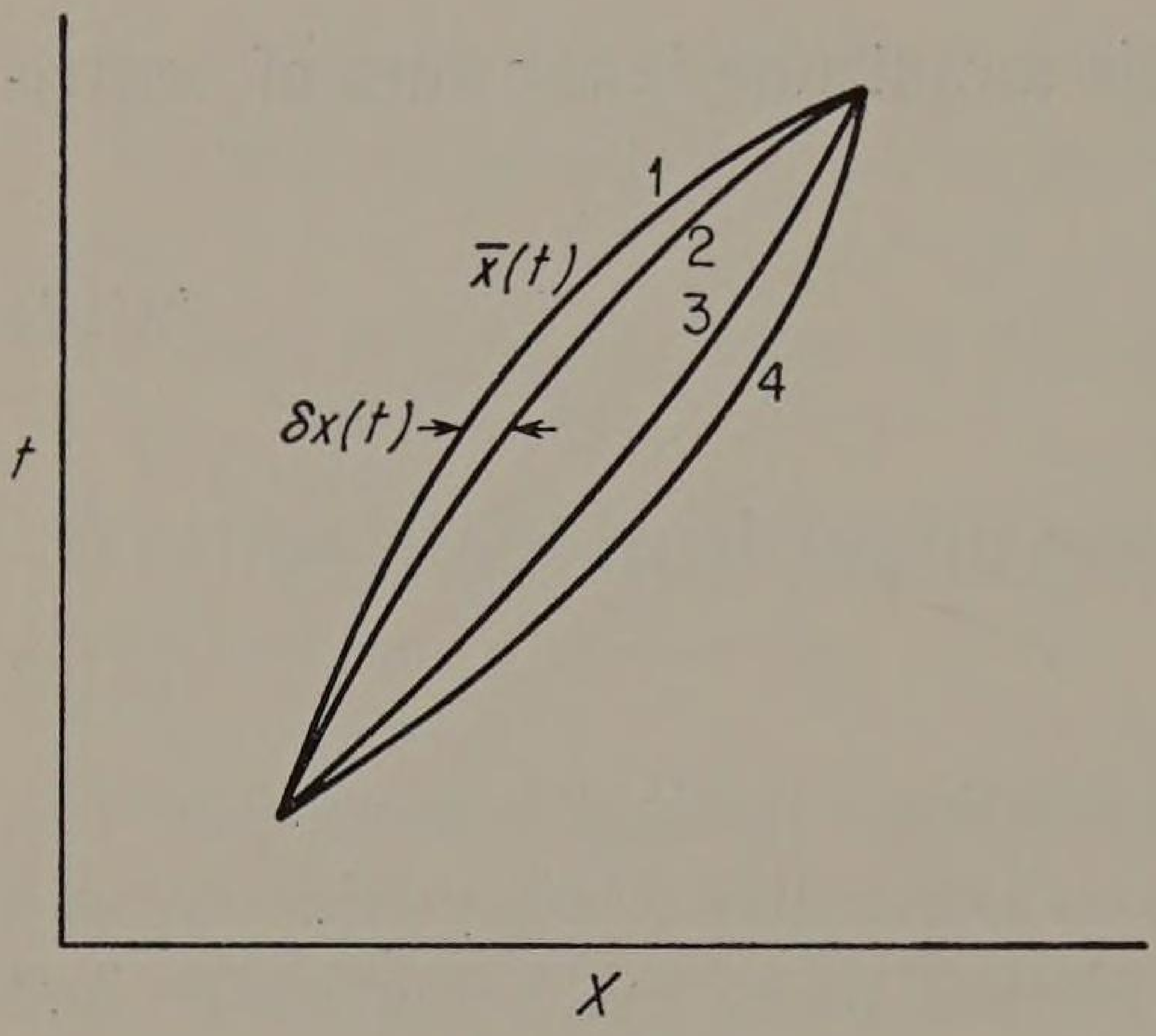


Fig. 2-1 The classical path 1, $\bar{x}(t)$, is that for which a certain integral, the action S, is minimum. If the path is varied by $\delta x(t)$, to path 2, the integral suffers no first-order change. This determines the equation of motion.

In quantum mechanics, the amplitude to go from a to b is the sum of amplitudes for each interfering alternative path. The amplitude for a given path, $e^{iS/\hbar}$, has a phase proportional to the action.

If the action is very large compared to $\hbar$, neighboring paths such as 3 and 4 have slightly different actions. Such paths will (because of the smallness of $\hbar$) have very different phases. Their contributions will cancel out. Only in the vicinity of the classical path $\bar{x}(t)$, where the action changes little when the path varies, will neighboring paths, such as 1 and 2, contribute in the same phase and constructively interfere. That is why the approximation of classical physics—that only the path $\bar{x}(t)$ need be considered—is valid when the action is very large compared to $\hbar$.

tory as being of importance. In this way the classical laws of motion arise from the quantum laws.

We may note that trajectories which differ from $\bar{x}$ contribute as long as the action is still within about $\hbar$ of S_{cl}. The classical trajectory is indefinite to this slight extent, and this rule serves as a measure of the limitations of the precision of the classically defined trajectory.

Next consider the dependence of the phase on the position of the end point (x_b,t_b). If we change the end point a little, this phase changes a great deal, and $K(b,a)$ changes very rapidly. If by a "smooth function" we mean one like S_{cl} which changes only when changes in argument which are appreciable on a classical scale are made, we note that $K(b,a)$ is far from smooth, but in this classical approximation our arguments show that it is of the form

$$K(b,a) = \text{“smooth function”} \cdot e^{(i/\hbar)S_{cl}} \tag{2-16}$$

All these approximate considerations apply to a situation on a scale for which classical physics might be expected to work ($S \gg \hbar$). But at an atomic level, S may be comparable with $\hbar$, and then all trajectories must be added in Eq. (2-14) in detail. No particular trajectory is of overwhelming importance, and of course Eq. (2-16) is not necessarily a good approximation. To deal with such cases, we shall have to find out how to carry out such sums as are implied by Eq. (2-14).

2-4 THE SUM OVER PATHS

Analogy with the Riemann Integral. Although the qualitative idea of a sum of a contribution for each of the paths is clear, a more precise mathematical definition of such a sum must be given. The number of paths is a high order of infinity, and it is not evident what measure is to be given to the space of paths. It is our purpose in this section to give such a mathematical definition. This definition will be found rather cumbersome for actual calculation. In the succeeding chapters we shall describe other and more efficient methods of computing the sum over all paths. As for this section, it is hoped that the mathematical difficulty, or rather inelegance, will not distract the reader from the physical content of the ideas.

We can begin our understanding with a consideration of the ordinary Riemann integral. We could say, very roughly, that the area A, under a curve, is the sum of all its ordinates. Better, we could say that it is proportional to that sum. But to make the idea precise, we do this: take a subset of all ordinates (e.g., those spaced at equal intervals h). Adding these ordinates, we obtain

$$A \sim \sum_i f(x_i) \tag{2-17}$$

where the summation is carried out over the finite set of points x_i, as shown in Fig. 2-2.

The next step is to define A as the limit of this sum as the subset of points, and thus the subset of ordinates, becomes more complete or—because a finite set is never any measurable part of the infinite continuum—we may better say as the subset becomes more representative of the complete set. We can pass to the limit in an orderly manner by taking continually smaller and smaller values of h. In so doing, we would obtain a different sum for each value of h. No limit exists. In order to obtain a limit to this process, we must specify some normalizing factor which should depend on h. Of course, for the Riemann integral, this factor is just h itself. Now the limit exists

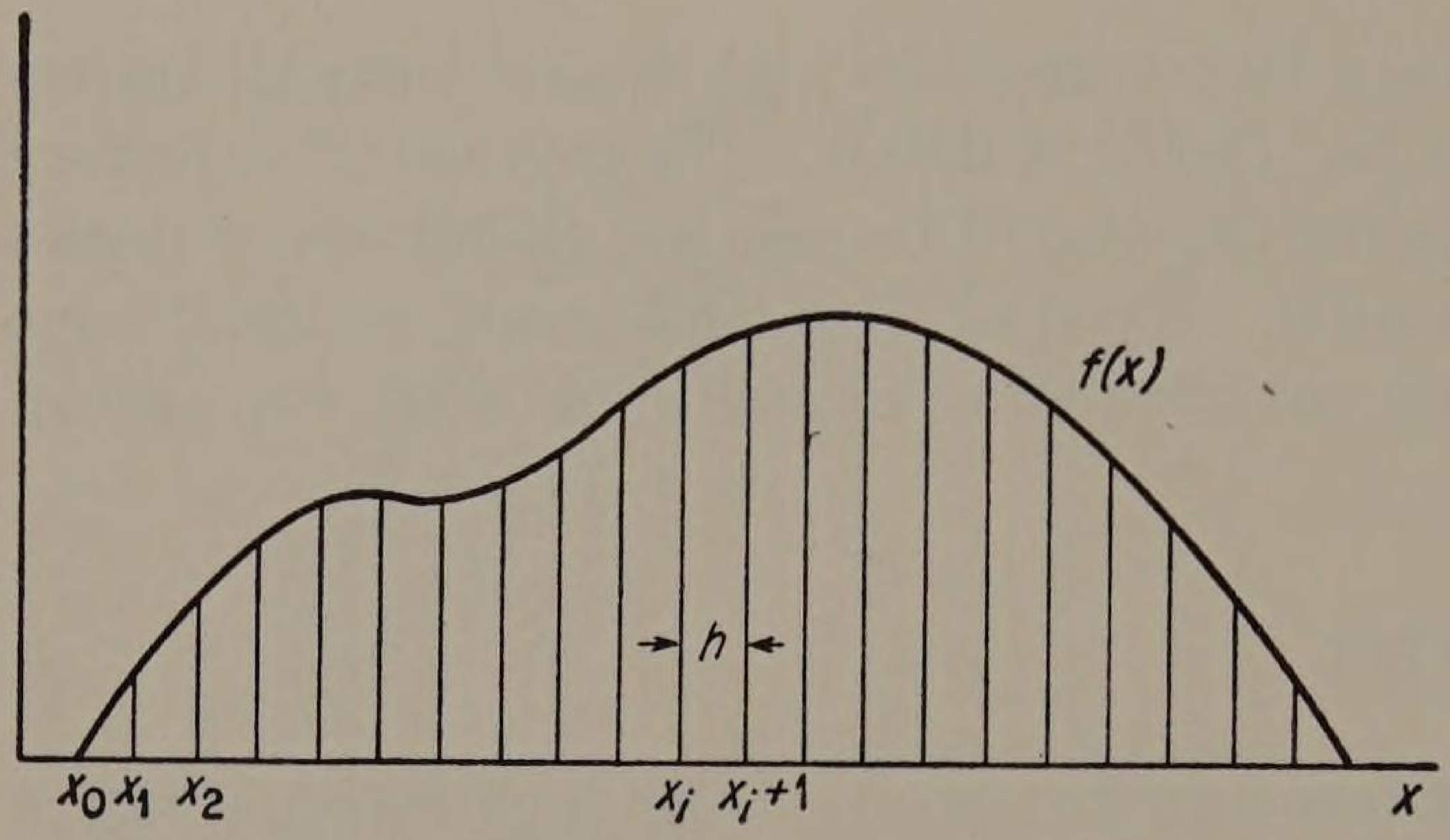


Fig. 2-2 In the definition of the ordinary Riemann integral, a set of ordinates is drawn from the abscissa to the curve. The ordinates are spaced a distance h apart. The integral (area between the curve and the abscissa) is approximated by h times the sum of the ordinates. This approximation approaches the correct value as h approaches zero.

An analogous definition can be used for path integrals. The measure which goes to zero in the limit process is the time interval ϵ between discrete points on the paths.

and we may write the expression

$$A = \lim \left[h \sum_i f(x_i) \right] \tag{2-18}$$

Constructing the Sum. We can follow through an analogous procedure in defining the sum over all paths. First, we choose a subset of all paths. To do this, we divide the independent variable time into steps of width ϵ. This gives us a set of values t_i spaced a distance ϵ apart between the values t_a and t_b. At each time t_i we select some special point x_i. We construct a path by connecting all the points so selected with straight lines. It is possible to define a sum over all paths constructed in this manner by taking a multiple integral over all values of x_i for i between 1 and $N - 1$, where

$$\begin{aligned} N\epsilon &= t_b - t_a \\ \epsilon &= t_{i+1} - t_i \\ t_0 &= t_a \qquad t_N = t_b \\ x_0 &= x_a \qquad x_N = x_b \end{aligned} \tag{2-19}$$

The resulting equation is

$$K(b,a) \sim \int\int \cdots \int \phi[x(t)]\, dx_1\, dx_2 \cdots dx_{N-1} \tag{2-20}$$

We do not integrate over x_0 or x_N because these are the fixed end points x_a and x_b. This equation corresponds formally to Eq. (2-17).

In the present case we can obtain a more representative sample of the complete set of all possible paths between a and b by making ϵ smaller. However, just as in the case of the Riemann integral, we cannot proceed to the limit of this process because the limit does not exist. Once again we must provide some normalizing factor which we expect will depend upon ϵ.

Unfortunately, to define such a normalizing factor seems to be a very difficult problem and we do not know how to do it in general terms. But we do know how to give the definition for all situations which so far seem to have practical value. For example, take the case where the lagrangian is given by Eq. (2-2). The normalizing factor turns out to be A^{-N}, where

$$A = \left(\frac{2\pi i\hbar\epsilon}{m}\right)^{1/2} \tag{2-21}$$

We shall see later (e.g., Sec. 4-1) how this result is obtained. With this factor the limit exists and we may write

$$K(b,a) = \lim_{\epsilon\to 0} \frac{1}{A} \int\int \cdots \int e^{(i/\hbar)S[b,a]} \frac{dx_1}{A}\frac{dx_2}{A} \cdots \frac{dx_{N-1}}{A} \tag{2-22}$$

where

$$S[b,a] = \int_{t_a}^{t_b} L(\dot{x},x,t)\,dt \tag{2-23}$$

is a line integral taken over the trajectory passing through the points x_i with straight sections in between, as in Fig. 2-3.

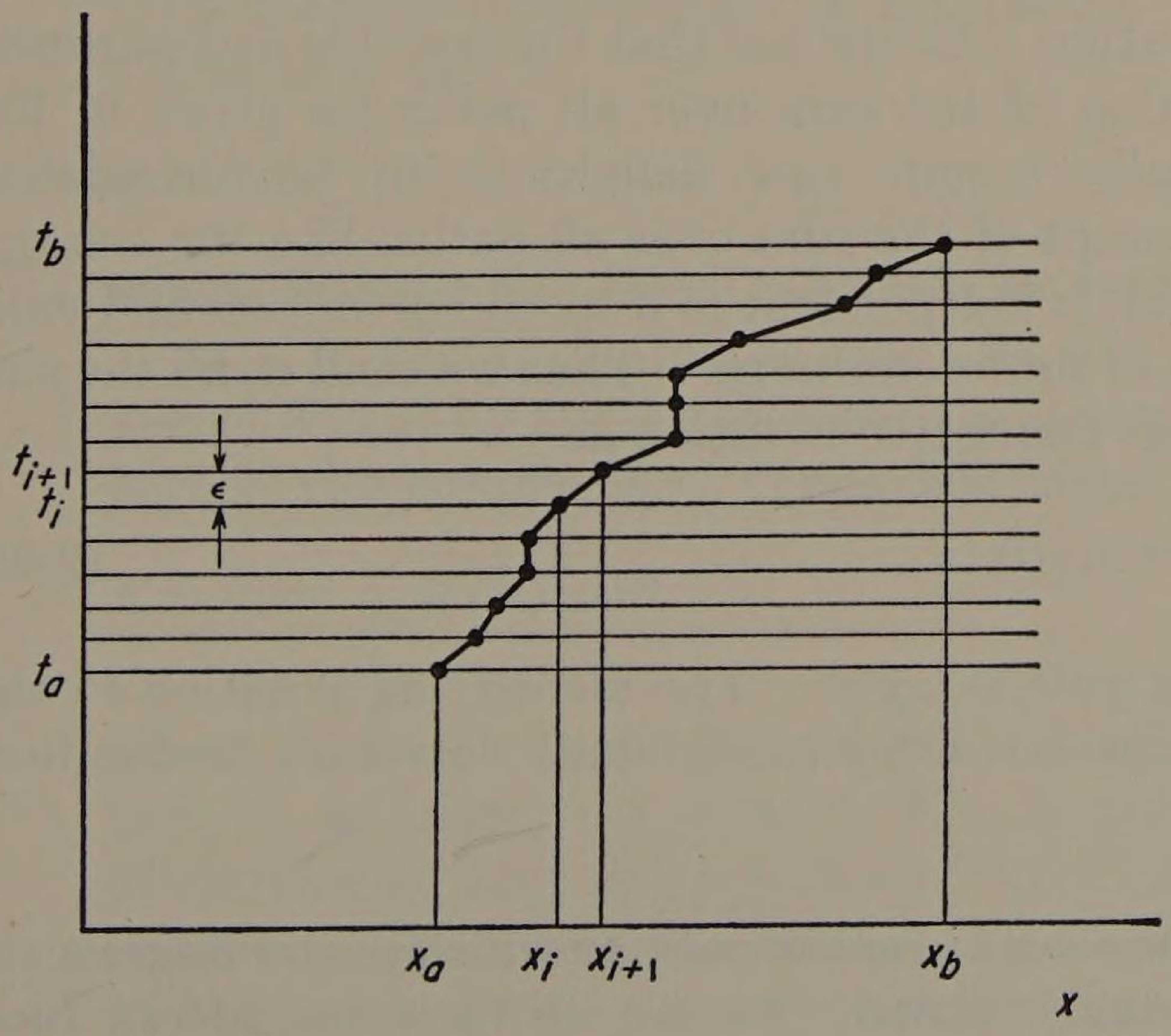


Fig. 2-3 The sum over paths is defined as a limit, in which at first the path is specified by giving only its coordinate x at a large number of specified times separated by very small intervals ϵ. The path sum is then an integral over all these specific coordinates. Then to achieve the correct measure, the limit is taken as ϵ approaches 0.

It is possible to define the path in a somewhat more elegant manner. Instead of straight lines between the points i and $i + 1$, we could use sections of the classical orbit. Then we could say that S is the minimum value of the integral of the lagrangian over all the paths which go through the specified points (x_i,t_i). With this definition no recourse is made to arbitrary straight lines.

The Path Integral. There are many ways to define a subset of all the paths between a and b. The particular definition we have used here may not be the best for some mathematical purposes. For example, suppose the lagrangian depends upon the acceleration of x. In the way we have constructed the path, the velocity is discontinuous at the various points (x_i,t_i); that is, the acceleration is infinite at these points. It is possible that this situation would lead to trouble. However, in the few such examples with which we have had experience the substitution

$$\ddot{x} = \frac{1}{\epsilon^2}\,(x_{i+1} - 2x_i + x_{i-1}) \tag{2-24}$$

has been adequate. There may be other cases where no such substitution is available or adequate, and the present definition of a sum over all paths is just too awkward to use. Such a situation arises in ordinary integration in which the Riemann definition, as in Eq. (2-18), is not adequate and recourse must be had to some other definition, such as that of the Lebesgue.

The necessity to redefine the method of integration does not destroy the concept of integration. So we feel that the possible awkwardness of the special definition of the sum over all paths [as given in Eq. (2-22)] may eventually require new definitions to be formulated. Nevertheless, the concept of the sum over all paths, like the concept of an ordinary integral, is independent of a special definition and valid in spite of the failure of such definitions. Thus we shall write the sum over all paths in a less restrictive notation as

$$K(b,a) = \int_a^b e^{(i/\hbar)S[b,a]}\,\mathfrak{D}x(t) \tag{2-25}$$

which we shall call a *path integral*. The identifying notation in this expression is the script $\mathfrak{D}$. Only rarely shall we return to the form given in Eq. (2-22).

Problem 2-6 The class of functionals for which path integrals can be defined is surprisingly varied. So far we have considered func-

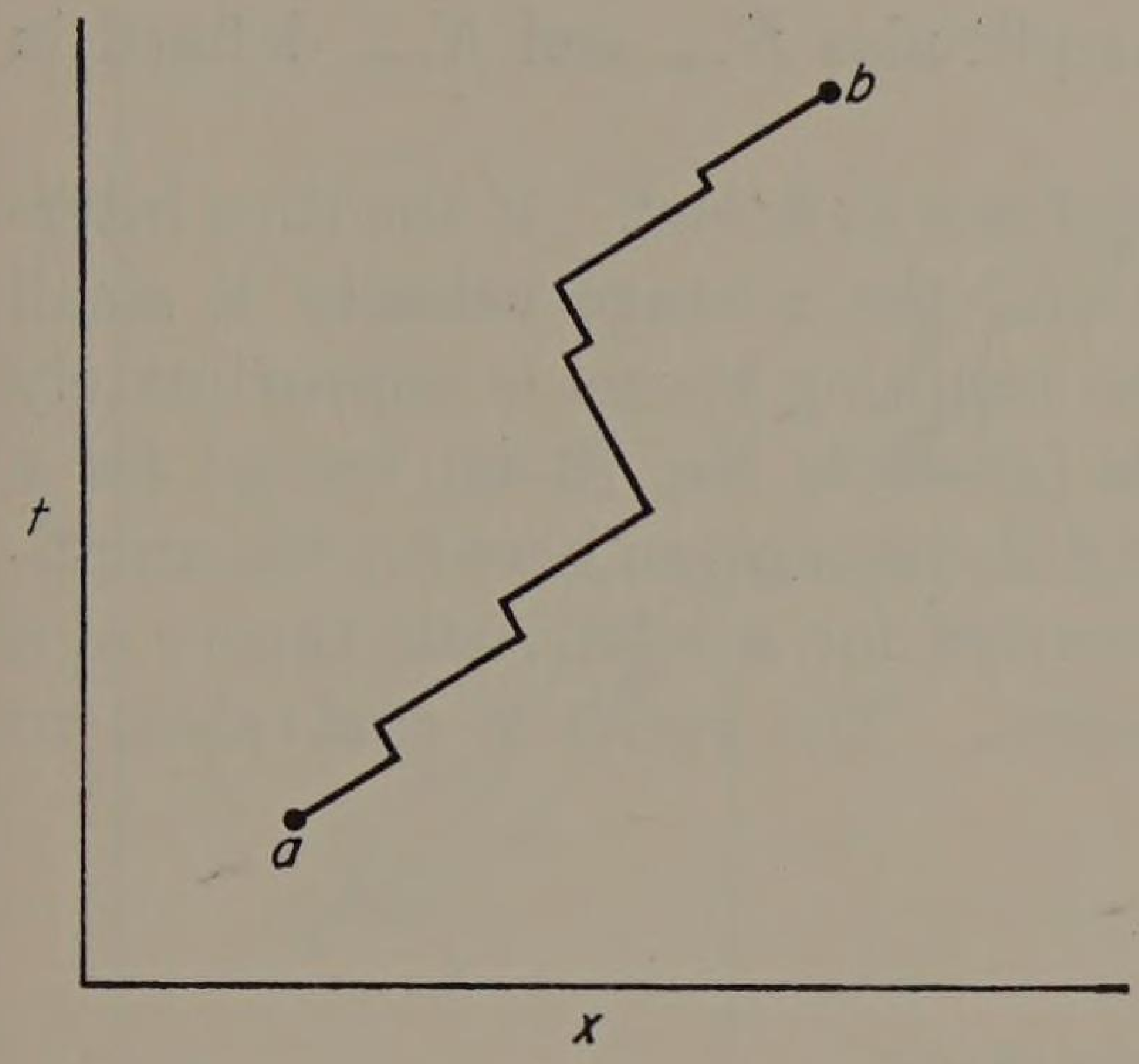


Fig. 2-4 The path of a relativistic particle traveling in two dimensions is a zigzag of straight segments. The slope of the segments is constant in magnitude and differs only in sign from zig to zag. The amplitude for a particular path, as well as the kernel to go from a to b, depends on the number of corners R along a path, as shown by Eqs. (2-26) and (2-27).

tionals such as that given in Eq. (2-15). Here we shall consider quite a different type. This latter type of functional arises in a one-dimensional relativistic problem. Suppose a particle moving in one dimension can go only forward or backward at the velocity of light. For convenience, we shall define the units such that the velocity of light, the mass of the particle, and Planck's constant are all unity. Then in the xt plane all trajectories shuttle back and forth with slopes of ±45°, as in Fig. 2-4. The amplitude for such a path can be defined as follows: Suppose time is divided into small equal steps of length ϵ. Suppose reversals of path direction can occur only at the boundaries of these steps, i.e., at $t = t_a + n\epsilon$, where n is an integer. For this relativistic problem the amplitude to go along such a path is different from the amplitude defined in Eq. (2-15). The correct definition for the present case is

$$\phi = (i\epsilon)^R \tag{2-26}$$

where R is the number of reversals, or corners, along the path.

As a problem, the reader may use this definition to calculate the kernel $K(b,a)$ by adding together the contribution for the paths of one corner, two corners, etc. Thus determine

$$K(b,a) = \sum_R N(R)(i\epsilon)^R \tag{2-27}$$

where $N(R)$ is the number of paths possible with R corners. It is best to calculate four separate K's, namely, the amplitude $K_{++}(b,a)$ of starting at the point a with a positive velocity and coming into the point b with a positive velocity, the amplitude $K_{+-}(b,a)$ of starting at the point a with a negative velocity and coming into the point b

with a positive velocity, and the amplitudes K_{-+} and K_{--} defined in a similar fashion.

Next suppose the unit of time is defined as $\hbar/mc^2$. If the time interval is very long ($t_b - t_a \gg \hbar/mc^2$) and the average velocity is small $[x_b - x_a \ll c(t_b - t_a)]$, show that the resulting kernel is approximately the same as that for a free particle [given in Eq. (3-3)], except for a factor $\exp[(imc^2/\hbar)(t_a - t_b)]$. The definition given here for the amplitude, and the resulting kernel, is correct for a relativistic theory of a free particle moving in one dimension. The result is equivalent to the Dirac equation for that case.

2-5 EVENTS OCCURRING IN SUCCESSION

The Rule for Two Events. In this section we shall derive an important law for the composition of amplitudes for events which occur successively in time. Suppose t_c is some time between t_a and t_b. Then the action along any path between a and b can be written as

$$S[b,a] = S[b,c] + S[c,a] \tag{2-28}$$

This follows from the definition of the action as an integral in time and also from the fact that L does not depend on derivatives higher than the velocity. (Otherwise, we would have to specify values of velocity and perhaps higher derivatives at point c.) Using Eq. (2-25) to define the kernel, we can write

$$K(b,a) = \int e^{(i/\hbar)S[b,c]+(i/\hbar)S[c,a]}\,\mathfrak{D}x(t) \tag{2-29}$$

It is possible to split any path into two parts. The first part would have the end points x_a and $x_c = x(t_c)$, and the second part would have the end points x_c and x_b, as shown in Fig. 2-5. It is possible to integrate over all paths from a to c, then over all paths from c to b, and finally integrate the result over all possible values of x_c. In performing the first step of this integration $S[b,c]$ is constant. Thus the result can be written as

$$K(b,a) = \int_{x_c}\int_c^b e^{(i/\hbar)S[b,c]}K(c,a)\,\mathfrak{D}x(t)\,dx_c \tag{2-30}$$

where integrations must now be carried out not only over paths between c and b but also over the variable end point x_c. In the next step we carry out the integration over all paths between some point with an arbitrary x_c and the point b. All that is left is an integral

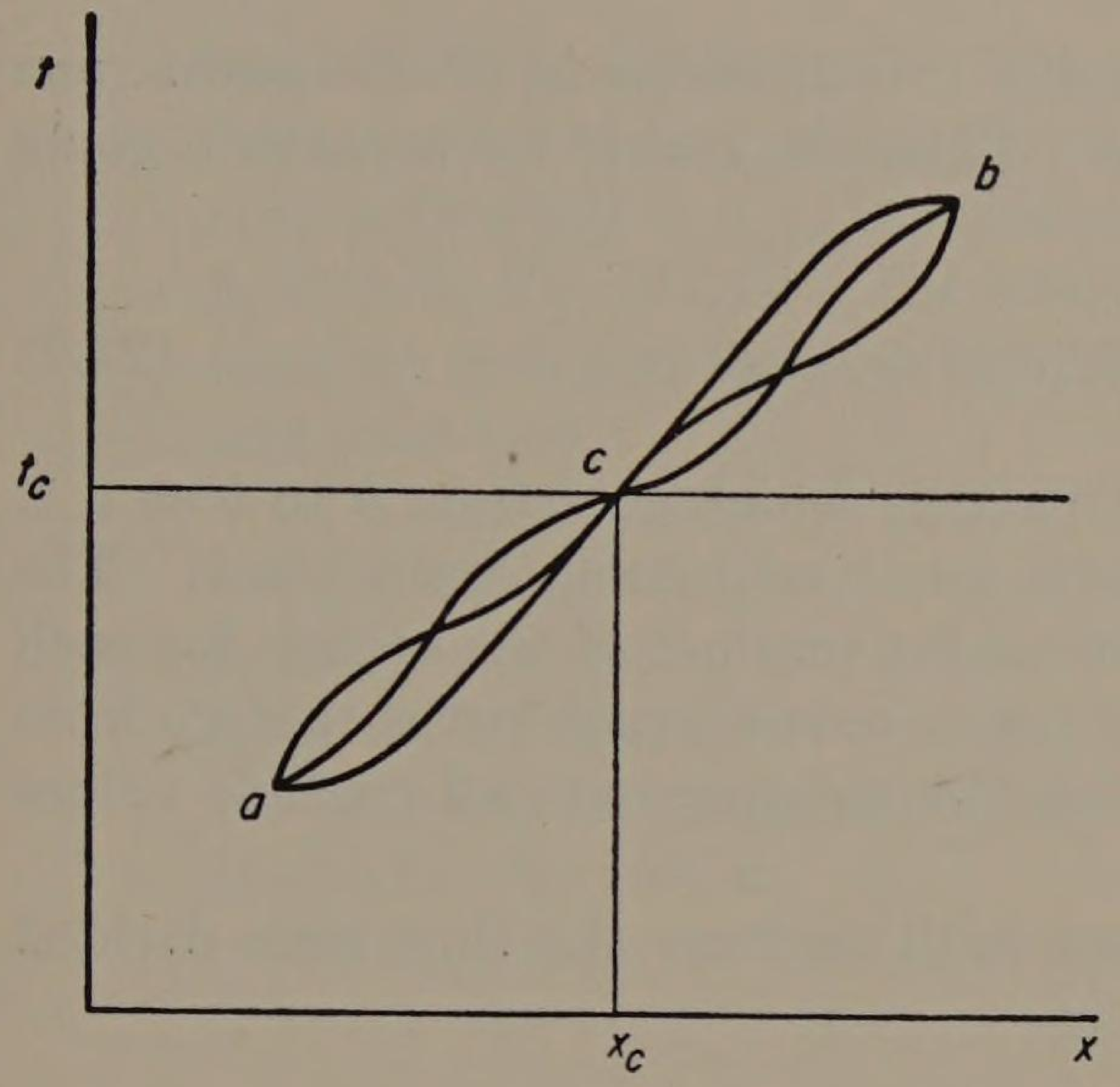


Fig. 2-5 One way the sum over all paths can be taken is by first summing over paths which go through the point at x_c and time t_c and later summing over the points x_c. The amplitude on each path that goes from a to b via c is a product of two factors: (1) an amplitude to go from a to c and (2) an amplitude to go from c to b. This is therefore valid also for the sum over all paths through c: the total amplitude to go from a to b via c is $K(b,c)K(c,a)$. Thus summing over the alternatives (values of x_c), we get for the total amplitude to go from a to b, Eq. (2-31).

over all possible values of x_c. Thus

$$K(b,a) = \int_{x_c} K(b,c)K(c,a)\,dx_c \tag{2-31}$$

Perhaps the argument is clearer starting from Eq. (2-22). Select one of the discrete times as t_c. Thus let $t_c = t_k$ and $x_c = x_k$. First carry out all the integrations over x_i, such that $i < k$. This will introduce the factor in the integral $K(c,a)$. Next carry out the integrals over all x_i, such that $i > k$. This introduces the factor $K(b,c)$. All that is left is an integral over x_c. The result can be written as Eq. (2-31).

This result can be summarized in the following way. All alternative paths from a to b can be labeled by specifying the position x_c through which they pass at the time t_c. Then the kernel for a particle going from a to b can be computed from the rules:

1. The kernel to go from a to b is the sum, over all possible values of x_c of amplitudes, for the particle to go from a to c and then to b.
2. The amplitude to go from a to c and then to b is the kernel to go from a to c times the kernel to go from c to b.

Thus we have the rule: *Amplitudes for events occurring in succession in time multiply.*

Extension to Several Events. There are many applications for this important rule, and several will be developed in succeeding chapters. Here we shall show the application wherein we follow an alternative route in deriving the equation for the kernel, Eq. (2-22).

It is perfectly possible to make two divisions in all the paths: one at t_c and the other at, say, t_d. Then the kernel for a particle going from a to b can be written as

$$K(b,a) = \int_{x_c}\int_{x_d} K(b,c)K(c,d)K(d,a)\,dx_c\,dx_d \tag{2-32}$$

This means that we look at a particle which goes from a to b as if it went first from a to d, then from d to c, and finally from c to b. The amplitude to follow such a path is the product of the kernels for each part of the path. The kernel taken over all such paths that go from a to b is obtained by integrating this product over all possible values of x_c and x_d.

We can continue this process until we have the time scale divided into N intervals. The result is

$$K(b,a) = \int_{x_1}\int_{x_2}\cdots\int_{x_{N-1}} K(b, N-1)K(N-1, N-2)\cdots K(i+1, i)\cdots K(1,a)\,dx_1\,dx_2\cdots dx_{N-1} \tag{2-33}$$

This means that we can define the kernel in a manner different from that given in Eq. (2-22). In this alternative definition the kernel for a particle to go between two points separated by an infinitesimal time interval ϵ is

$$K(i+1, i) = \frac{1}{A}\exp\left[\frac{i\epsilon}{\hbar}L\left(\frac{x_{i+1}-x_i}{\epsilon}, \frac{x_{i+1}+x_i}{2}, \frac{t_{i+1}+t_i}{2}\right)\right] \tag{2-34}$$

which is correct to first order in ϵ. Then by the rules for multiplying amplitude of events which occur successively in time, we have

$$\phi[x(t)] = \lim_{\epsilon\to 0}\prod_{i=0}^{N-1} K(i+1, i) \tag{2-35}$$

for the amplitude of a complete path. Then, using the rule that amplitudes for alternative paths add, we arrive at a definition for $K(b,a)$. It can be seen that the resulting expression is actually the same as Eq. (2-22).

2-6 SOME REMARKS

In the relativistic theory of the electron we shall not find it possible to express the amplitude for a path as $e^{iS/\hbar}$, or in any other simple way. However, the laws for combining amplitudes still work (with some small modifications). The amplitude for a trajectory still exists. As a matter of fact, it is still given by Eq. (2-35). The only difference is that $K(i + 1, i)$ is not so easily expressed in a relativistic theory as it is in Eq. (2-34). The complications arise from the necessity to consider spin and the possibility of the production of pairs of electrons and positrons.

In nonrelativistic systems with a larger number of variables, and even in the quantum theory of electromagnetic field, not only do the laws for combining amplitudes still hold but the amplitude itself follows the rules set down in this chapter. That is, each motion of a variable has an amplitude whose phase is $1/\hbar$ times the action associated with the motion.

We shall take up these more complicated examples in later chapters.

3

Developing the Concepts with Special Examples

In this chapter we shall develop the kernels governing some special types of motion. We shall explore the physical meaning of the mathematical results in order to develop some physical intuition about motion under quantum-mechanical laws. The wave function will be introduced and its relation to the kernel will be described. This represents the first step in connecting the present approach to quantum mechanics with the more traditional approaches.

We shall also introduce some special mathematical methods for computing the sum over all paths. The idea of a sum over all paths was described in Chap. 2 with the help of a particular computational method. Although that method may clarify the concept, it is an awkward tool with which to work. The simpler methods to be introduced in this chapter will be of great use in our future work.

Thus the present chapter has three purposes: deepening our understanding of quantum-mechanical principles, beginning the connection between our present approach and other approaches, and introducing some useful mathematical methods.

3-1 THE FREE PARTICLE

The Path Integral. The method used in Chap. 2 to describe a sum over all paths will be used here to compute the kernel for a free particle. The lagrangian for a free particle is

$$L = m\frac{\dot{x}^2}{2} \tag{3-1}$$

Thus with the help of Eqs. (2-21) to (2-23) the kernel for a free particle is

$$K(b,a) = \lim_{\epsilon\to 0} \iint \cdots \int \exp\left[\frac{im}{2\hbar\epsilon}\sum_{i=1}^{N}(x_i - x_{i-1})^2\right] dx_1 \cdots dx_{N-1}\left(\frac{2\pi i\hbar\epsilon}{m}\right)^{-N/2} \tag{3-2}$$

This represents a set of gaussian integrals, i.e., integrals of the form $\int[\exp(-ax^2)]\,dx$ or $\int[\exp(-ax^2 + bx)]\,dx$. Since the integral of a gaussian is again a gaussian, we may carry out the integrations on one variable after the other. After the integrations are completed, the limit may be taken. The result is

$$K(b,a) = \left[\frac{2\pi i\hbar(t_b - t_a)}{m}\right]^{-1/2} \exp\frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)} \tag{3-3}$$

The calculation is carried out as follows. Notice first that

$$\int_{-\infty}^{\infty}\left(\frac{2\pi i\hbar\epsilon}{m}\right)^{-2/2}\exp\left\{\frac{m}{2i\hbar\epsilon}[(x_2-x_1)^2+(x_1-x_0)^2]\right\}dx_1$$
$$=\left(\frac{2\pi i\hbar\cdot 2\epsilon}{m}\right)^{-1/2}\exp\left[\frac{m}{2i\hbar(2\epsilon)}(x_2-x_0)^2\right]\quad(3\text{-}4)$$

Next we multiply this result by

$$\left(\frac{2\pi i\hbar\epsilon}{m}\right)^{-1/2}\exp\left[\frac{m}{2i\hbar\epsilon}(x_3-x_2)^2\right]\quad(3\text{-}5)$$

and integrate again, this time over x_2. The result is similar to that of Eq. (3-4), except that $(x_2 - x_0)^2$ becomes $(x_3 - x_0)^2$ and the expression 2ϵ is replaced by 3ϵ in two places. Thus we get

$$\left(\frac{2\pi i\hbar\cdot 3\epsilon}{m}\right)^{-1/2}\exp\left[\frac{m}{2i\hbar\cdot 3\epsilon}(x_3-x_0)^2\right]$$

In this way a recursion process is established which, after $n - 1$ steps, gives

$$\left(\frac{2\pi i\hbar n\epsilon}{m}\right)^{-1/2}\exp\left[\frac{m}{2i\hbar\cdot n\epsilon}(x_n-x_0)^2\right]$$

Since $n\epsilon = t_n - t_0$, it is easy to see that the result after $N - 1$ steps is identical with Eq. (3-3).

There is an alternative procedure. Equation (3-4) can be used to integrate over all the variables x_i for which i is odd (assuming N is even). The result is an expression formally like Eq. (3-2), but with half as many variables of integration. The remaining variables are defined at points in time spaced a distance 2ϵ apart. Hence, at least in the case that N is of the form 2^k, Eq. (3-3) results from k steps of this kind.

Problem 3-1 The probability that a particle arrives at the point b is by definition proportional to the absolute square of the kernel $K(b,a)$. For the free-particle kernel of Eq. (3-3) this is

$$P(b)\,dx=\frac{m}{2\pi\hbar(t_b-t_a)}\,dx\quad(3\text{-}6)$$

Clearly this is a relative probability, since the integral over the complete range of x diverges. What does this particular normalization mean? Show that this corresponds to a classical picture in which a particle starts from the point a with all momenta equally likely. Show

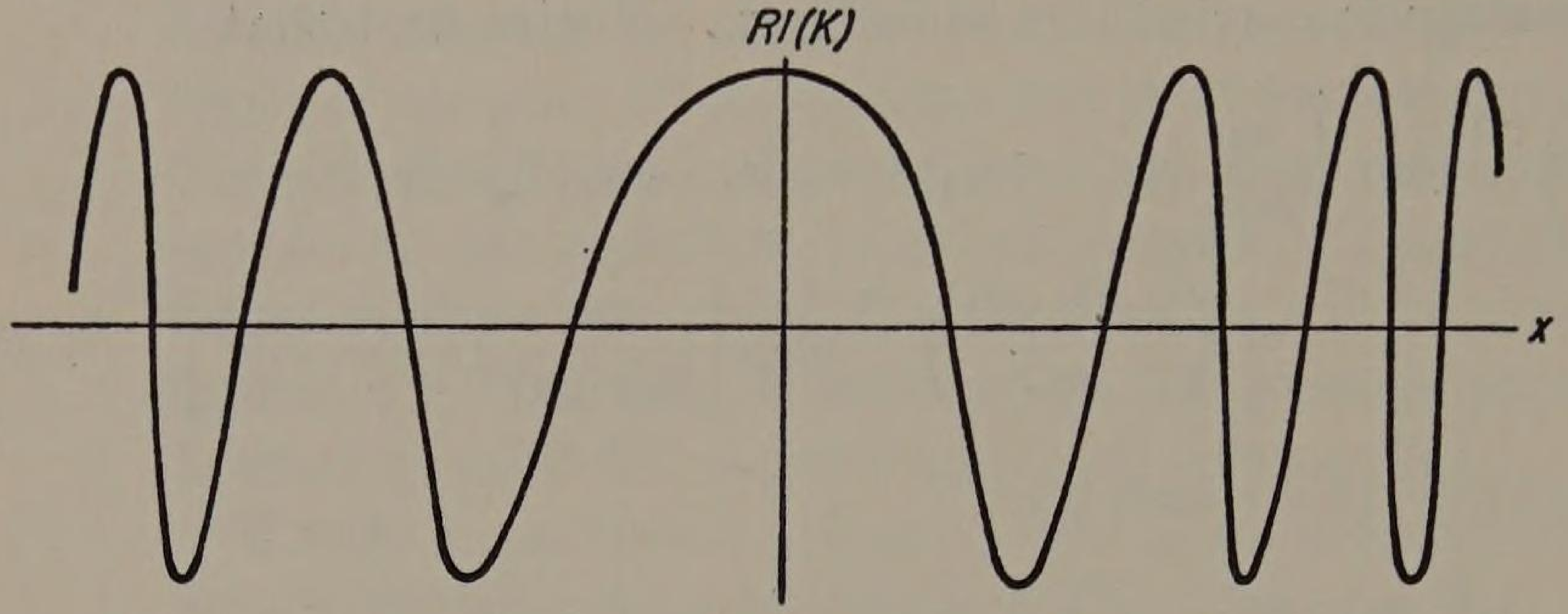


Fig. 3-1 The real part of amplitude to arrive at various distances x from the origin after a time t. The imaginary part (not shown) is an analogous wave 90° out of phase, so that the absolute square of the amplitude is constant. The wavelength is short at large x, namely, where a classical particle could arrive only if it moved with high velocity. Generally, the wavelength and classical momentum are inversely related [Eq. (3-10)].

that the corresponding relative probability that the momentum of the particle lies in the range dp is $dp/2\pi\hbar$.

Momentum and Energy. We now study some of the implications of the free-particle kernel. For convenience let the point a represent the origin in both time and space. The amplitude to go to some other point $b = (x,t)$ is

$$K(x,t;0,0) = \left(\frac{2\pi i\hbar t}{m}\right)^{-1/2} \exp\frac{imx^2}{2\hbar t} \tag{3-7}$$

If time is fixed, the amplitude varies with distance as shown in Fig. 3-1, in which the real part of Eq. (3-7) is plotted.

We see that as we get farther from the origin the oscillations become more and more rapid. If x is so large that many oscillations have occurred, then the distance between successive nodes is nearly constant, at least for the next few oscillations. That is, the amplitude behaves much like a sine wave of slowly varying wavelength. It is interesting to evaluate this wavelength λ. Changing x by λ must increase the phase of the amplitude by 2π. That is,

$$2\pi = \frac{m(x+\lambda)^2}{2\hbar t} - \frac{mx^2}{2\hbar t} = \frac{mx\lambda}{\hbar t} + \frac{m\lambda^2}{2\hbar t} \tag{3-8}$$

Neglecting the quantity λ^2 relative to $x\lambda$ (that is, assuming $x \gg \lambda$), we find

$$\lambda = \frac{2\pi\hbar}{m(x/t)} \tag{3-9}$$

From a classical point of view a particle which moves from the origin to x in the time interval t has a velocity x/t and a momentum mx/t. From the quantum-mechanical point of view, when the motion can be adequately described by assigning a classical momentum to the particle of $p = mx/t$, then the amplitude varies in space with the wavelength

$$\lambda = \frac{h}{p} \tag{3-10}$$

We may show this relation still more generally. Suppose we have some large piece of apparatus, such as a magnetic analyzer, which is supposed to bring particles of a given momentum p to a given point. We shall show that, whenever the apparatus is large enough that classical physics offers a good approximation, then the amplitude for a particle to arrive at the prescribed point varies rapidly in space with a wavelength equal to h/p. For as we have seen, in such a situation, the kernel is approximated by

$$K \sim \exp\left[\frac{i}{\hbar} S_{cl}(b,a)\right] \tag{3-11}$$

Changes in the position of the final point x_b cause variations in the classical action. If this action is large compared to $\hbar$ (the semiclassical approximation), the kernel will oscillate very rapidly with changes in x_b. The change in phase per unit displacement of the end point is

$$k = \frac{1}{\hbar}\frac{\partial S_{cl}}{\partial x_b} \tag{3-12}$$

but $\partial S_{cl}/\partial x_b$ is the classical momentum of the particle when it arrives at the point x_b (see Prob. 2-4). Thus $p = \hbar k$. This quantity k, the phase change per unit distance of a wave, is called the wave number, and it is very convenient to use. Since the wavelength is the distance over which the phase changes by 2π, then $k = 2\pi/\lambda$. Equation (3-12) is de Broglie's formula relating the momentum to the wave number of a wave, $p = \hbar k$.

Next, let us study the time dependence of the free-particle kernel given by Eq. (3-7). Suppose we hold the distance fixed and vary the time. The variation of the real part of the kernel is shown in Fig. 3-2. Both frequency and amplitude change with t.

Suppose t is very large and neglect the change in amplitude with variations of t. The period of oscillation T is defined as the time

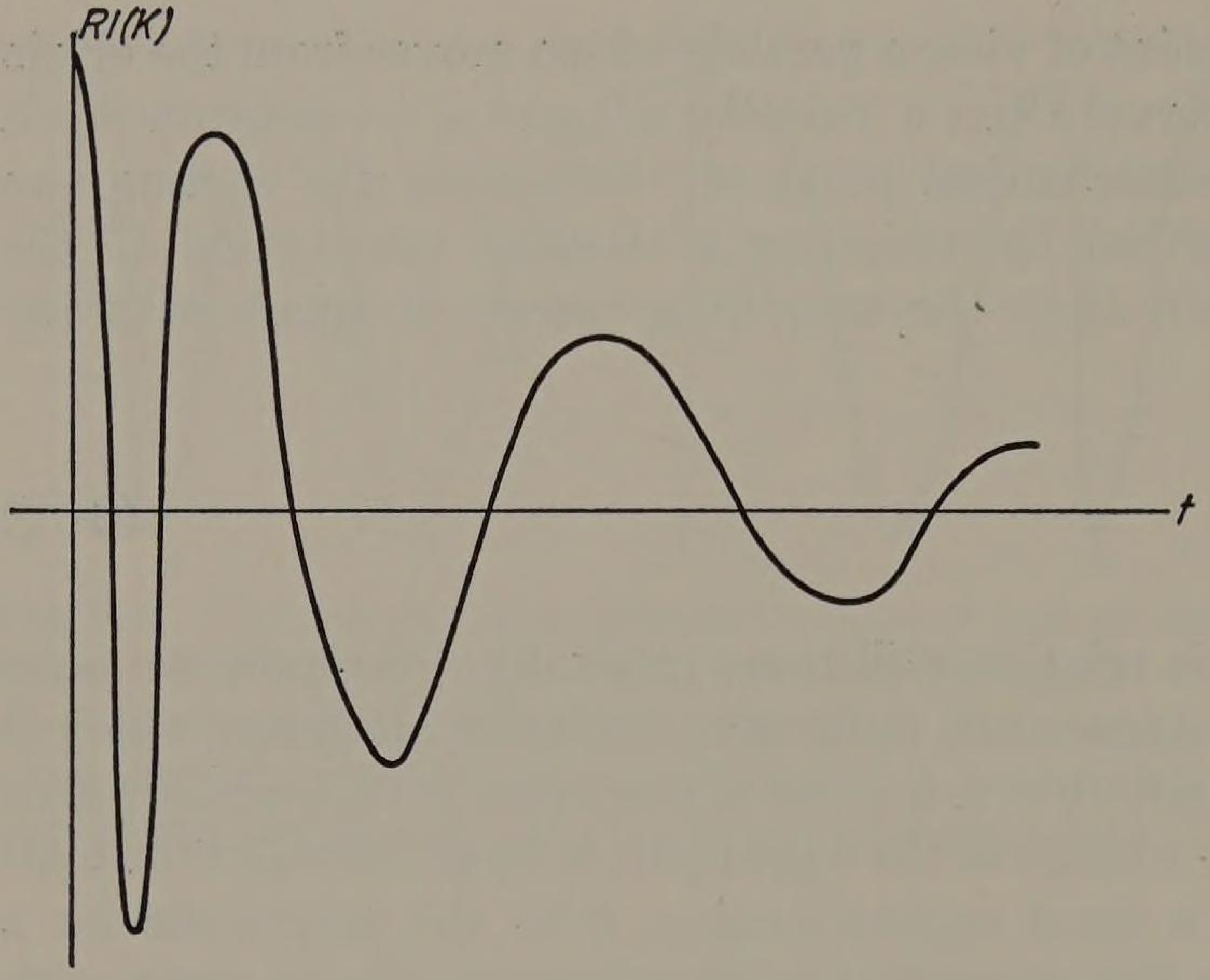


Fig. 3-2 The amplitude to find the particle at a given point varies with time. The real part is plotted here. The frequency of the oscillations is proportional to the energy that a classical particle would have to have to arrive at the point in question within the time interval t.

required to increase the phase by 2π. Thus

$$2\pi = \frac{mx^2}{2\hbar t} - \frac{mx^2}{2\hbar(t+T)} = \frac{mx^2}{2\hbar t^2}\left(\frac{T}{1+T/t}\right) \tag{3-13}$$

By introducing the angular frequency $\omega = 2\pi/T$, and assuming $t \gg T$, we can write this equation as

$$\omega \simeq \frac{m}{2\hbar}\left(\frac{x}{t}\right)^2 \tag{3-14}$$

Since $m(x/t)^2/2$ is the classical energy of a free particle, this equation says

$$\text{Energy} = \hbar\omega \tag{3-15}$$

This relation, like the one relating momentum and wavelength, holds for any apparatus which can be adequately described by classical physics; and, like the previous relation, it can be obtained from a more general argument.

Referring to Eq. (3-11), any variation of the time t_b of an end point will cause a rapid oscillation of the kernel. The resulting frequency is

$$\omega = \frac{1}{\hbar}\frac{\partial S_{cl}}{\partial t} \tag{3-16}$$

The quantity $\partial S_{cl}/\partial t$ is interpreted classically as the energy E (refer to Prob. 2-5). Thus

$$\omega = \frac{E}{\hbar} \tag{3-17}$$

In this way the concepts of momentum and energy are extended to quantum mechanics by the following rules:

1. If the amplitude varies as e^{ikx}, we say that the particle has momentum $\hbar k$.
2. If the amplitude has a definite frequency varying in time as $e^{-i\omega t}$, we say the energy is $\hbar\omega$.

We have just shown that this rule will agree with the definition of energy and momentum in the classical limit.

Problem 3-2 Show by substitution that the free-particle kernel $K(b,a)$ satisfies the differential equation

$$-\frac{\hbar}{i}\frac{\partial K(b,a)}{\partial t_b} = -\frac{\hbar^2}{2m}\frac{\partial^2 K(b,a)}{\partial x_b^2} \tag{3-18}$$

whenever t_b is greater than t_a.

3-2 DIFFRACTION THROUGH A SLIT

The Conceptual Experiment. We can learn more about the physical interpretation of quantum mechanics and its relation to classical mechanics by considering another somewhat more complicated example. Suppose a particle is liberated at $t = 0$ from the origin and then, at an interval of time T later, we observe that it is at a certain point x_0. Classically, we would say that the particle has had a velocity $v_0 = x_0/T$. The implication would be that if a particle were left alone to continue for another interval of time τ, it would move an additional distance $v_0\tau$. To analyze this quantum-mechanistically, we shall attempt to solve the following problem:

At $t = 0$ the particle starts from the origin $x = 0$. After an interval T we shall suppose that it is known that the particle is within the distance $\pm b$ of x_0. We ask: After an additional interval τ, what is the probability of finding the particle at an additional displacement x from the position x_0? The net amplitude to arrive at the position x at the time $t + \tau$ can be considered as the sum of contributions from every trajectory that goes from the origin to the final point, provided that that trajectory lies in the interval $\pm b$ from x_0 at the time T.

We shall calculate this in a moment, but first it is worth remarking on what kind of an experiment we are contemplating here. How can we know that the particle passes the point x_0 within the interval $\pm b$? One way would be to make an observation of the particle at the time

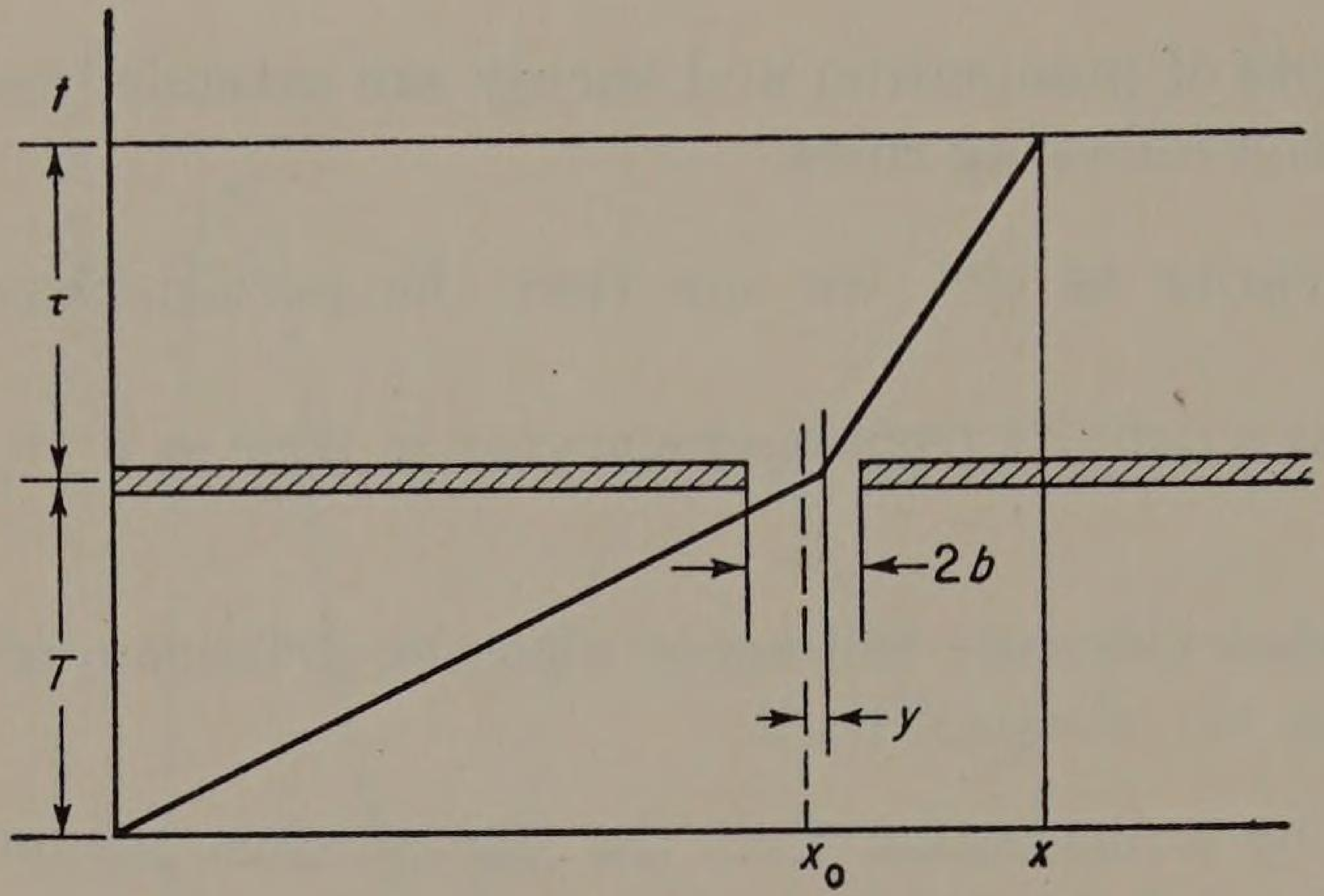


Fig. 3-3 A particle starting at $x = 0$ when $t = 0$ is determined to pass between $x_0 - b$ and $x_0 + b$ at $t = T$. We wish to calculate the probability of finding the particle at some point x at a time τ sec later, i.e., when $t = T + \tau$. According to classical laws, the particle would have to be between $x_0(\tau/T) + b(1 + \tau/T)$ and $x_0(\tau/T) - b(1 + \tau/T)$, that is, between the rectilinear extensions of the original slit. However, quantum-mechanical laws show that such particles have a nonzero probability of appearing outside these classical limits.

We cannot approach this problem by a single application of the free-particle law of motion, since the particle is actually constrained by the slit. So we break the problem up into two successive free-particle motions. The first takes the particle from $x = 0$ at $t = 0$ to $x = x_0 + y$ at $t = T$, where $|y| \leq b$. The second takes the particle from $x_0 + y$ at $t = T$ to x at $t = T + \tau$. The overall amplitude is an integral of the product of these two free-particle kernels, as shown by Eq. (3-19).

T to see if it is within the interval $\pm b$. This would be the most natural way to proceed, but it is somewhat more difficult to analyze in detail (because of the complicated interaction between the electron and the observing mechanism) than another way of doing the experiment.

Suppose we look, say, with very strong light, everywhere all along the x axis *except* within $\pm b$ from the point x_0 at the time T. If we find the particle, we discontinue the experiment. We consider only those cases in which a thorough investigation of the region, *except* for the region $\pm b$, shows no particle is present. That is, all trajectories which pass outside the limits $\pm b$ from x_0 are rejected. The experimental situation is diagramed in Fig. 3-3. The amplitude then can be written as

$$\psi(x) = \int_{-b}^{b} K(x + x_0, T + \tau; x_0 + y, T)K(x_0 + y, T; 0, 0)\, dy \qquad (3\text{-}19)$$

This expression is written down in accordance with the rule for combining amplitudes of events occurring in succession in time. The first event is that the particle goes from the origin to the slit. The second event is that the particle proceeds from the slit to the point x further on. The slit has a finite width, and passage through each elemental interval of the slit represents an alternative way of proceeding along the complete path. Thus we must integrate over the width of the slit. All particles which miss the slit are captured and removed from the experiment. Amplitudes for such particles are not included. All the particles which get through the slit move as free particles with kernels given by Eq. (3-3). Thus the amplitude is

$$\psi(x) = \int_{-b}^{b} \left(\frac{2\pi i\hbar\tau}{m}\right)^{-\frac{1}{2}} \left\{\exp\left[\frac{im(x-y)^2}{2\hbar}\right]\right\} \left(\frac{2\pi i\hbar T}{m}\right)^{-\frac{1}{2}} \left\{\exp\left[\frac{im(x_0+y)^2}{2\hbar T}\right]\right\} dy \qquad (3\text{-}20)$$

This integral can be expressed in terms of Fresnel integrals. Such a representation contains the physical results we are after, but in an obscure way owing to the mathematical complexity of the Fresnel integral form. Rather than confuse the physical results by mathematical complexity, we shall set up a different, but analogous, expression which leads to a simpler mathematical form.

The Gaussian Slit. Suppose we introduce a function $G(y)$ as a factor in the integrand. If this function is defined as unity between $y = -b$ and $y = +b$ and zero elsewhere, the limits of integration can be extended to infinity without any change in the result. Then

$$\psi(x) = \int_{-\infty}^{\infty} \frac{mG(y)}{2\pi i\hbar\sqrt{\tau T}} \left(\exp\left\{\frac{im}{2\hbar}\left[\frac{(x-y)^2}{\tau} + \frac{(x_0+y)^2}{T}\right]\right\}\right) dy \qquad (3\text{-}21)$$

where

$$\begin{aligned} G(y) &= 1 \qquad \text{for } -b \le y \le b \\ &= 0 \qquad \text{for } |y| > b \end{aligned}$$

Instead of this, suppose we define $G(y)$ to be a gaussian function, thus:

$$G(y) = e^{-y^2/2b^2} \qquad (3\text{-}22)$$

This function has the shape shown in Fig. 3-4. The *effective width* of such a curve is related to the parameter b. For this particular function, approximately two-thirds of the area under the curve lies between $-b$ and $+b$.

We do not know how to design metal parts for our imaginary experiment which will produce such a gaussian slit. However, there is no

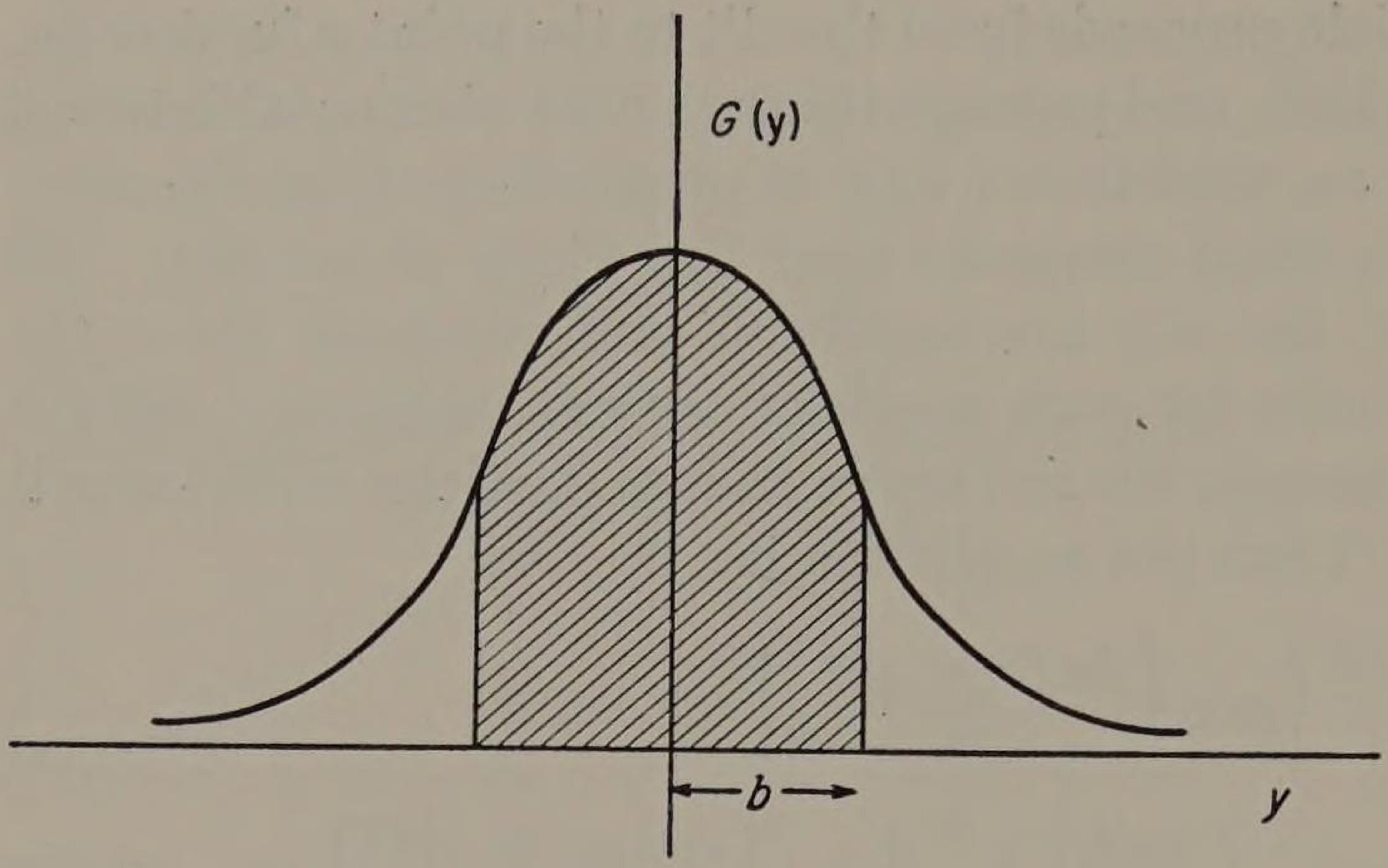


Fig. 3-4 The form of the gaussian function $G(y) = e^{-y^2/2b^2}$. The curve has the same shape as the normal distribution with a standard deviation b.

conceptual difficulty. We now have a situation in which the particles at time T are distributed along the x axis with a relative amplitude proportional to the function $G(y)$. (The relative probability is proportional to $[G(y)]^2$.) If the particles move classically, we would expect, after a succeeding interval of time τ, to find them similarly distributed along the x axis with a new center a distance x_1 beyond x_0 and an increased width parameter b_1 given by

$$x_1 = \left(\frac{x_0}{T}\right)\tau \qquad b_1 = b\left(1 + \frac{\tau}{T}\right) \tag{3-23}$$

as shown in Fig. 3-5.

With such a gaussian slit the equation for the amplitude is

$$\begin{aligned}\psi(x) = \int_{-\infty}^{\infty} \frac{m}{2\pi i\hbar\sqrt{\tau T}} &\left\{\exp\left[\frac{im}{2\hbar}\left(\frac{x^2}{\tau} + \frac{x_0{}^2}{T}\right)\right.\right.\\ &\left.\left. + \frac{im}{\hbar}\left(-\frac{x}{\tau} + \frac{x_0}{T}\right)y + \left(\frac{im}{2\hbar\tau} + \frac{im}{2\hbar T} - \frac{1}{2b^2}\right)y^2\right]\right\} dy\end{aligned} \tag{3-24}$$

This integral is of the form

$$\int_{-\infty}^{\infty} [\exp(\alpha x^2 + \beta x)]\,dx = \sqrt{\frac{\pi}{-\alpha}}\exp -\frac{\beta^2}{4\alpha} \qquad \text{for Re}(\alpha) \leq 0 \tag{3-25}$$

which is integrated by completing the square in the exponent. Thus the amplitude becomes

$$\begin{aligned}\psi(x) = \sqrt{\frac{m}{2\pi i\hbar}}\left[T\tau\left(\frac{1}{T} + \frac{1}{\tau} + \frac{\hbar i}{b^2m}\right)\right]^{-1/2} &\exp\left[\frac{im}{2\hbar}\left(\frac{x^2}{\tau} + \frac{x_0{}^2}{T}\right)\right.\\ &\left. - \frac{(im/\hbar)^2(-x/\tau + x_0/T)^2}{4(im/2\hbar)(1/\tau + 1/T + \hbar i/b^2m)}\right]\end{aligned} \tag{3-26}$$

The classical velocity to get from the origin to the center of the slit is $v_0 = x_0/T$. When we use this as a substitution and rearrange some of the terms, the expression for the amplitude becomes

$$\psi(x) = \sqrt{\frac{m}{2\pi i\hbar}} \left(T + \tau + T\tau \frac{\hbar i}{mb^2}\right)^{-1/2}$$
$$\exp\left[\frac{im}{2\hbar}\left(v_0{}^2T + \frac{x^2}{\tau}\right) + \frac{(m^2/2\hbar^2\tau^2)(x - v_0\tau)}{(m/\hbar)(i/T + i/\tau) - 1/b^2}\right] \tag{3-27}$$

We shall consider first the relative probability for the particle to arrive at various points along the x axis. This probability is proportional to the absolute square of the amplitude. The absolute value of an exponent with an imaginary argument is 1. So, by rationalizing the second factor and the denominator in the last exponent of Eq. (3-27), we obtain

$$P(x)\,dx = \frac{m}{2\pi\hbar T}\,\frac{b}{\Delta x} \exp\left[\frac{-(x - v_0\tau)^2}{(\Delta x)^2}\right] dx \tag{3-28}$$

where we have used the substitution

$$(\Delta x)^2 = b^2\left(1 + \frac{\tau}{T}\right) + \frac{\tau^2\hbar^2}{m^2b^2}$$
$$= b_1{}^2 + \frac{\tau^2\hbar^2}{m^2b^2} \tag{3-29}$$

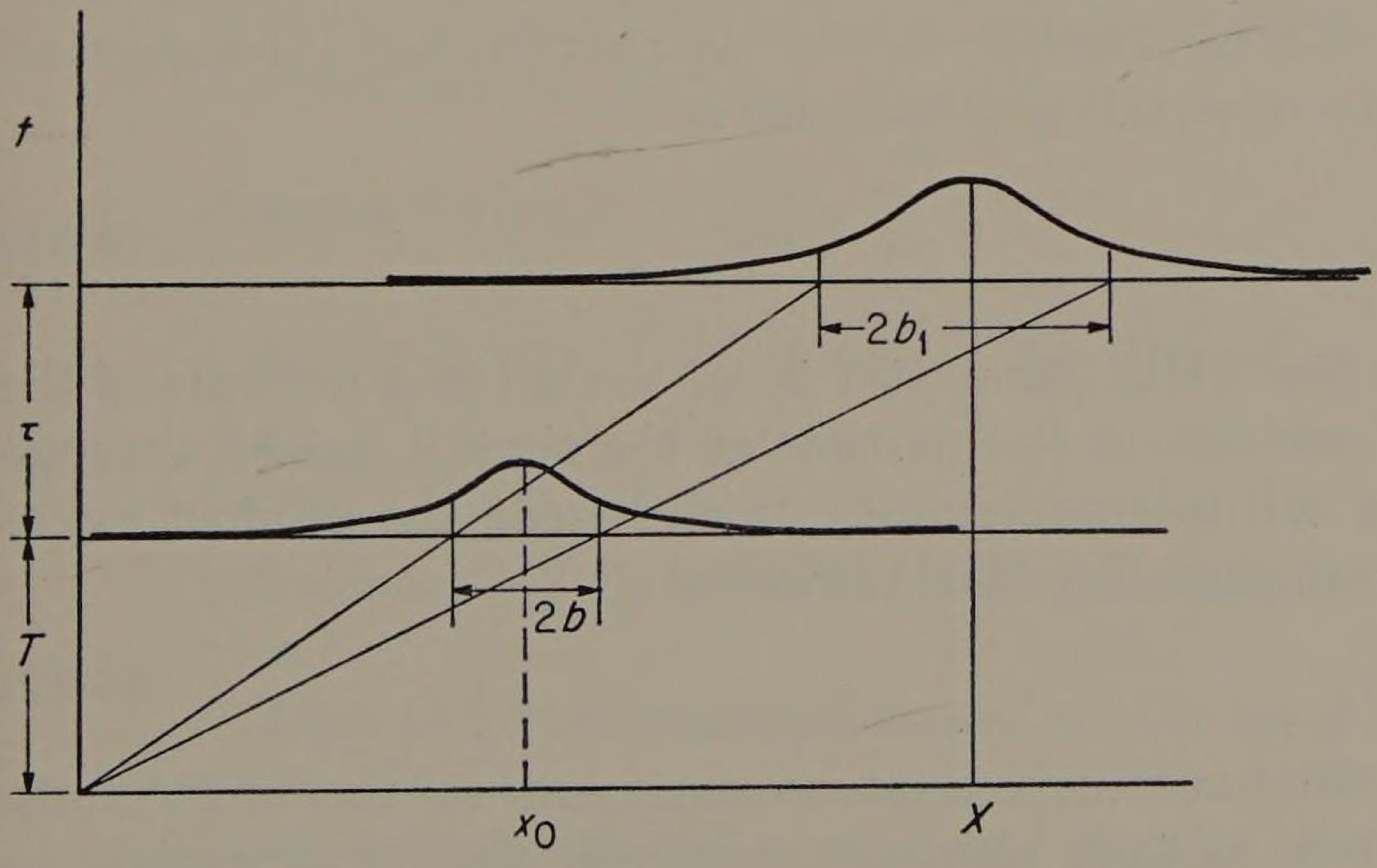


Fig. 3-5 The paths of particles moving through a gaussian slit. If the particles obeyed classical laws of motion, then the distribution of particles at time $T + \tau$ would have the same form as the distribution at time T. The difference would be only a spreading out proportional to the time of flight. The characteristic width of the distribution would be increased from $2b$ to $2b_1$, where $b_1 = b(T + \tau)/T$. For quantum-mechanical motion, the actual spreading is greater than this.

As we expected, the distribution is a gaussian centered about the point $x_1 = v_0\tau$ of Eq. (3-23). However, the spread of the distribution, Δx, is larger than the expected value b_1 of Eq. (3-23). This can be interpreted in the following manner. Suppose a_1 and a_2 are two independent random variables whose root-mean-square deviations about their average values are respectively α_1 and α_2. Then if $a_3 = a_1 + a_2$, the rms deviation of a_3 about its average value is $\alpha_3 = (\alpha_1^2 + \alpha_2^2)^{1/2}$. Now, the rms deviation in a particular distribution is a measure of the spread, or width, of the distribution. As a matter of fact, for the gaussian distribution $e^{-x^2/2b^2}$ the rms value is b.

Thus in the present case we find that the quantum-mechanical system acts as if it had an extra random variable x_1 whose rms deviation is

$$\Delta x_1 = \frac{\hbar\tau}{mb} \tag{3-30}$$

It is this extra deviation Δx_1, or spreading, rather than the apparent extra variable x_1, which has physical significance. That this term is quantum-mechanical in nature is clear from the inclusion of the constant $\hbar$. Such a term is important for particles of small mass and for narrow slits.

Thus quantum mechanics tells us that for small particles, passage through a narrow slit makes the future position uncertain. This uncertainty Δx_1 is proportional to the time interval τ between passage through the slit and the next observation of position. If we introduce the classical notion of velocity, we can say that passage through a slit causes a velocity uncertainty whose size is

$$\delta v = \frac{\hbar}{mb} \tag{3-31}$$

We could take the width parameter $2b$ of the slit as a measure of the uncertainty of position of the particle at the time it passed through the slit. If we call this uncertainty δx and write the product mv as the momentum p, then Eq. (3-31) becomes

$$\delta p\, \delta x = 2\hbar \tag{3-32}$$

Once more we have arrived at a statement of the uncertainty principle. It tells us that, although classically the velocity might be known, the future position has an additional uncertainty as though a random momentum had been generated by passing through a slit of width δx. If classical concepts are used to describe the results of quantum mechanics qualitatively, then we would say that knowledge of position creates uncertainty in momentum.

What about the factors that appear in front of the exponent in Eq. (3-28)? If we integrate this expression over the complete range of x from $-\infty$ to $+\infty$, the result is

$$P(\text{any } x) = \frac{m}{2\pi\hbar T} b\sqrt{\pi} \tag{3-33}$$

This must be the probability that the particle gets through the slit, since the integration includes those particles and only those particles which did get through. But we have another way of obtaining this result. Suppose we take the absolute square of the kernel $K(x_0 + y, T; 0, 0)$, which comprises the second half of the integrand in Eq. (3-20). This is just the probability per unit distance that the particle arrives at the point $x_0 + y$ in the slit. That is,

$$P(x_0 + y)\, dy = \frac{m}{2\pi\hbar T}\, dy \tag{3-34}$$

This probability is independent of the position along the slit. Thus, if we were to multiply it by the width of this slit, we would obtain the total probability for the particle to arrive at the slit. This implies that the effective width of our gaussian slit is $\sqrt{\pi}\, b$. Had we used the original sharp-edged slit, we would find the effective width to be $2b$.

Problem 3-3 By squaring the amplitude given in Eq. (3-20) and then integrating over x, show that the probability of passage through the original slit is

$$P(\text{going through}) = \frac{m}{2\pi\hbar T} 2b \tag{3-35}$$

In the course of this problem the integral

$$\int_{-\infty}^{\infty} e^{iax}\, dx \tag{3-36}$$

will appear. This is the integral representation of the Dirac δ function of a.[†]

Thus the quantum-mechanical results agree with the idea that the probability that a particle goes through a slit is equal to the probability that the particle arrives at the slit.

Momentum and Energy. Next we shall verify again that, when the momentum is definite, the amplitude varies as e^{ikx}. We return to a detailed study of the amplitude given in Eq. (3-26). This time we shall try to arrange conditions in our experiment so that the particle velocity after passing through the slit is known as accurately as possible.

[†] See the table of integrals in the Appendix and L. I. Schiff, "Quantum Mechanics," 2d ed., pp. 50–52, McGraw-Hill Book Company, New York, 1955.

Quite apart from any quantum-mechanical considerations, there is a classical uncertainty of b/T in the velocity. For any given slit width we can make this uncertainty negligible by choosing T very large. We can also make x_0 extremely large so that the average velocity $x_0/T = v_0$ does not go to 0. Considering v_0 and τ as constants, the expression for the amplitude in the limit of large T is

$$\psi(x) \simeq \frac{\text{const}}{\sqrt{1 + \tau\hbar/2mb^2}} \exp\left[\frac{imx^2}{2\hbar\tau} + \frac{m^2(x - \tau v_0)^2}{4\hbar^2\tau^2(im/2\hbar\tau - 1/2b^2)}\right] \tag{3-37}$$

Next we must arrange that the quantum-mechanical uncertainty in momentum $\hbar/b$ is very small. That is, we take b very large, so we can neglect $1/b^2$. Then the amplitude can be written as

$$\psi(x) \simeq \text{const} \exp\left(\frac{imv_0}{\hbar}x - \frac{imv_0{}^2}{2\hbar}\tau\right) \tag{3-38}$$

This is an important result. It says that, if we have arranged things so that the momentum of a particle is known to be p, then the amplitude for the particle to arrive at the point x at the time t is

$$\psi(x) \simeq \text{const} \exp\left(\frac{i}{\hbar}px - \frac{i}{\hbar}\frac{p^2}{2m}t\right) \tag{3-39}$$

We notice that this is a wave of definite wave number $k = p/\hbar$. Furthermore, it has a definite frequency $\omega = p^2/2m\hbar$. This means we can say that a free particle of momentum p has a definite quantum-mechanical energy (defined as $\hbar$ times frequency) which is $p^2/2m$ just as in classical mechanics.

The probability of arriving at any particular x, which is proportional to the square of the amplitude, is in this case independent of x. Thus exact knowledge of velocity means no knowledge of position. In arranging the experiment to give an accurately known velocity we have lost our chances for an accurate prediction of position. We have already seen that the reverse is true. The existence of the quantum-mechanical spread, inversely proportional to the slit with $2b$, implies that an exact knowledge of position precludes any knowledge of velocity. So, if you know where it is, you cannot say how fast it is going; and, if you know how fast it is going, you cannot say where it is. This again illustrates the uncertainty principle.

3-3 RESULTS FOR A SHARP-EDGED SLIT

Leaving the limiting case, suppose we return to a situation in which the slit width and quantum-mechanical spread are comparable in size and the times and distances of travel are not extremely large. We have seen that a gaussian slit leads to a gaussian distribution. If we use the more realistic sharp-edged version and work out the resulting Fresnel integrals, the probability distribution at the time τ after passing through the slit is like that shown in Fig. 3-6.

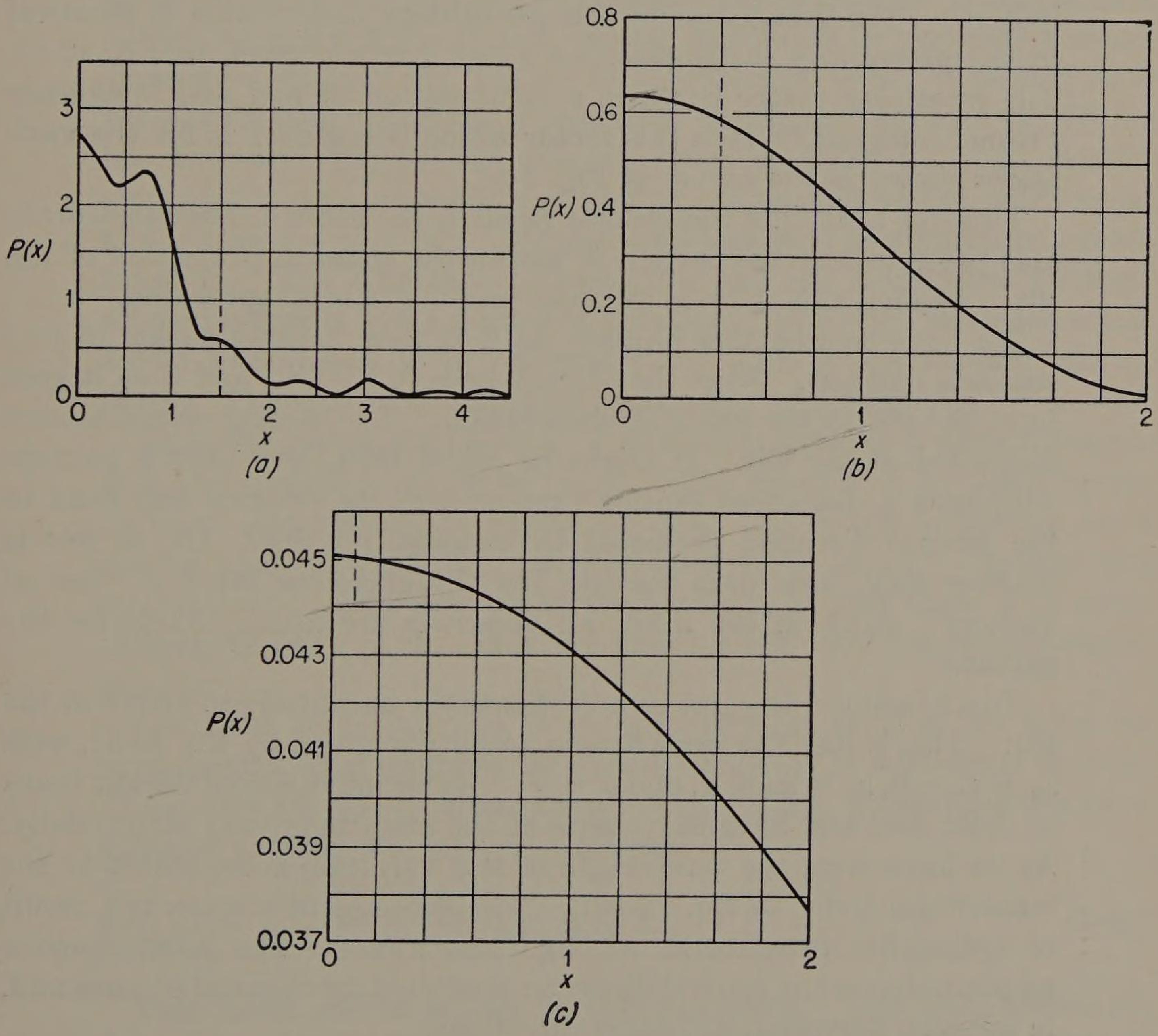


Fig. 3-6 The three curves show the distribution of electrons which have passed through sharp-edged slits of various widths. In each case the position of the classically predicted width $b_1 = b(1 + \tau/T)$ is shown by the dashed vertical line. Three different values have been chosen for the ratio of the classical width b_1 to the quantum-mechanical spreading Δx_1: curve (*a*), $b_1/\Delta x_1 = 15$; curve (*b*), $b_1/\Delta x_1 = 1$; and curve (*c*), $b_1/\Delta x_1 = 1/15$. In each case, the distribution spreads beyond the classical width. The rms width of the distribution is roughly proportional to $\Delta x = [(\Delta x_1)^2 + (b_1)^2]^{1/2}$.
]

This distribution is expressed by

$$P(x)\,dx = \frac{m}{2\pi\hbar(\tau + T)}\,\frac{1}{2}\,[C(u_1) - C(u_2)]^2 + \frac{1}{2}\,[S(u_1) - S(u_2)]^2\,dx \qquad (3\text{-}40)$$

where

$$u_1 = \frac{x - \tau v_0 - b(1 + \tau/T)}{(\pi\hbar\tau/m)(1 + \tau/T)} \qquad u_2 = \frac{x - \tau v_0 + b(1 + \tau/T)}{(\pi\hbar\tau/m)(1 + \tau/T)} \qquad (3\text{-}41)$$

and $C(u)$ and $S(u)$ are the real and imaginary parts of the Fresnel integral. The first factor in this probability distribution is identical to the probability distribution of a free particle given in Eq. (2-6). The remaining factor contains a combination of real and imaginary Fresnel integrals.[1] It is this factor which is responsible for the variations shown in the curves of Fig. 3-6.

Thus for both slits the general result is the same. The most probable place to find the particle is within the classical projection of the slit. Beyond this there is the quantum-mechanical spreading.

We have treated this problem as if it were a combination of two separate motions. First the particle goes to the slit, and then it goes from the slit to the point of observation. The motion seems almost disjointed at the slit. It might be asked then, how does a particle with such a disjointed motion "remember" its velocity and head in the general direction predicted by classical physics? Or, to put it another way, how does making the slit narrower cause a "loss of memory" until, in the limit, all velocities are equally likely for the particle?

To understand this, let us investigate the amplitude to arrive at the slit. This is just the free-particle amplitude given by Eq. (3-3), with $x_a = t_a = 0$, $x_b = x_0 + y$, and $t_b = T$. As we move across the slit (vary y), both real and imaginary parts of the amplitude vary sinusoidally. As we have seen, the wavelength of this variation is connected to the momentum [refer to Eq. (3-10)]. The subsequent motion is a result of optical-like interference among these waves. The interference is constructive in the general direction predicted by classical physics and, in general, destructive in other directions.

If there are many wavelengths across the slit (i.e., the slit is very wide) the resulting interference pattern is quite sharp and the motion is approximately classical. But suppose the slit is made so narrow that not even one whole wavelength is included. There are no longer any oscillations to give an interference, and velocity information is lost. Thus in the limit as the slit width goes to zero all velocities are equally likely for the particle.

---

[1] Refer to p. 34 of E. Jahnke and F. Emde, "Tables of Functions," Dover Publications, Inc., New York, 1943.

3-4 THE WAVE FUNCTION

We have developed the amplitude for a particle to reach a particular point in space and time by closely following its motion in getting there. However, it is often useful to consider the amplitude to arrive at a particular place without any special discussion of previous motion. Thus we would say $\psi(x,t)$ is the total amplitude to arrive at (x,t) from the past in some (perhaps unspecified) situation. Such an amplitude has the same probability characteristics as those we have already studied; i.e., the probability of finding the particle at the point x and at the time t is $|\psi(x,t)|^2$. We shall call this kind of amplitude a wave function. The difference between this and the amplitudes we have studied before is just a matter of notation. One often hears the statement: The system is in the "state" ψ. This is just another way of saying: The system is described by the wave function ψ.

Thus the kernel $K(x_2,t_2;x_1,t_1) = \psi(x_2,t_2)$ is actually a wave function. It is the amplitude to get to (x_2,t_2). The notation $K(x_2,t_2;x_1,t_1)$ gives us more information, in particular, that this is the amplitude for a special case in which the particle came from (x_1,t_1). Perhaps this information is of no interest to the problem, so that there is no point in keeping track of it. Then we just use the wave function notation $\psi(x_2,t_2)$.

Since the wave function is an amplitude, it satisfies the rules for combination of amplitudes for events occurring in succession in time. Thus since Eq. (2-31) is true for all points (x_1,t_1), we see that the wave function satisfies the integral equation

$$\psi(x_2,t_2) = \int_{-\infty}^{\infty} K(x_2,t_2;x_3,t_3)\psi(x_3,t_3)\,dx_3 \tag{3-42}$$

This result can be stated in physical terms. The total amplitude to arrive at (x_2,t_2) [that is, $\psi(x_2,t_2)$] is the sum, or integral, over all possible values of x_3 of the total amplitude to arrive at the point (x_3,t_3) [that is, $\psi(x_3,t_3)$] multiplied by the amplitude to go from 3 to 2 [that is, $K(x_2,t_2;x_3,t_3)$]. This means that the effects of all the past history of a particle can be expressed in terms of a single function. If we forget everything we knew about a particle except its wave function at a particular time, then we can calculate everything that can happen to that particle after that time. All of history's effect upon the future of the universe could be obtained from a single gigantic wave function.

Problem 3-4 Suppose a free particle has a definite momentum at the time $t = 0$ (that is, the wave function is $Ce^{ipx/\hbar}$). With the help of Eqs. (3-3) and (3-42), show that at some later time the particle has the same definite momentum (i.e., the wave function depends on x through the function $e^{ipx/\hbar}$) and varies in time as $e^{(-ip^2/2m\hbar)t}$. This means that the particle has the definite energy $p^2/(2m)$.

Problem 3-5 Use the results of Prob. 3-2 and Eq. (3-42) to show that the wave function satisfies the equation

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} \tag{3-43}$$

which is the Schrödinger equation for a free particle.

3-5 GAUSSIAN INTEGRALS

We are finished with the physical portion of this chapter, and we now proceed to mathematical considerations. We shall introduce some additional mathematical techniques which will help us to compute the sum over paths in certain situations.

The simplest path integrals are those in which all of the variables appear up to the second degree in an exponent. We shall call them gaussian integrals. In quantum mechanics this corresponds to a case in which the action S involves the path $x(t)$ up to and including the second power.

To illustrate how the method works in such a case, consider a particle whose lagrangian has the form

$$L = a(t)\dot{x}^2 + b(t)\dot{x}x + c(t)x^2 + d(t)\dot{x} + e(t)x + f(t) \tag{3-44}$$

The action is the integral of this function with respect to time between two fixed end points. Actually, in this form the lagrangian is a little more general than necessary. The factor $\dot{x}$ could be removed from those terms in which it is linear through an integration by parts, but this fact is unimportant for our present purpose. We wish to determine

$$K(b,a) = \int_a^b \exp\left[\frac{i}{\hbar}\int_{t_a}^{t_b} L(\dot{x},x,t)\right]\mathcal{D}x(t) \tag{3-45}$$

the integral over all paths which go from (x_a,t_a) to (x_b,t_b).

Of course, it is possible to carry out this integral over all paths in the way which was first described by dividing the region into short time elements, and so on. That this will work follows from the fact that the integrand is an exponential of a quadratic form in the variables $\dot{x}$ and x. Such integrals can always be carried out. But we shall not go through this tedious calculation, since we can determine the most important characteristics of the kernel in the following manner.

Let $\bar{x}(t)$ be the classical path between the specified end points. This is the path which is an extremum for the action S. In the notation we have been using

$$S_{cl}[b,a] = S[\bar{x}(t)] \tag{3-46}$$

We can represent x in terms of $\bar{x}$ and a new variable y:

$$x = \bar{x} + y \tag{3-47}$$

That is to say, instead of defining a point on the path by its distance $x(t)$ from an arbitrary coordinate axis, we measure instead the deviation $y(t)$ from the classical path, as shown in Fig. 3-7.

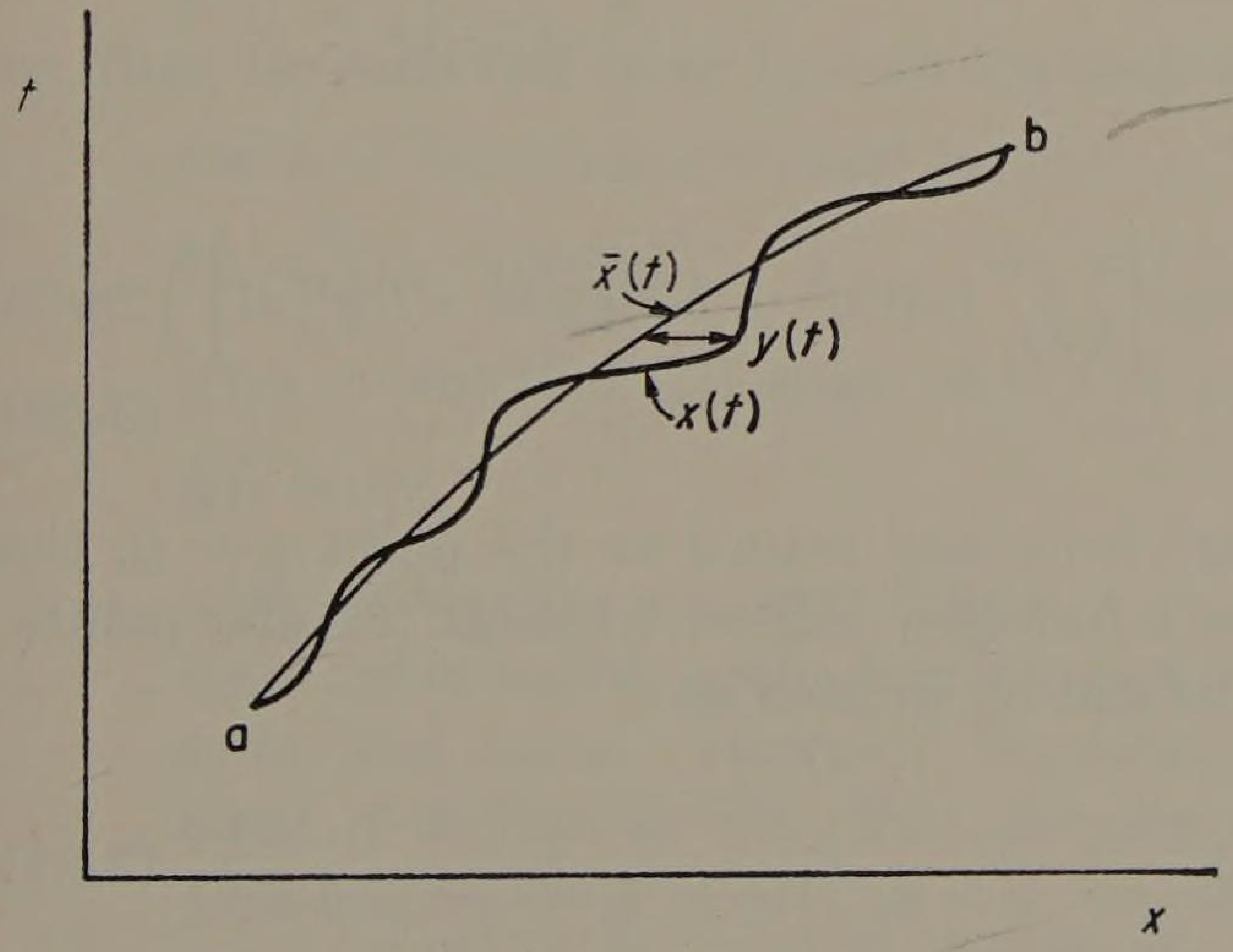


Fig. 3-7 The difference between the classical path $\bar{x}(t)$ and some possible alternative path $x(t)$ is the function $y(t)$. Since the paths must both reach the same end points, $y(t_a) = y(t_b) = 0$. In between these end points $y(t)$ can take any form. Since the classical path is completely fixed, any variation in the alternative path $x(t)$ is equivalent to the associated variation in the difference $y(t)$. Thus, in a path integral, the path differential $\mathfrak{D}x(t)$ can be replaced by $\mathfrak{D}y(t)$, and the path $x(t)$ by $\bar{x}(t) + y(t)$. In this form $\bar{x}(t)$ is a constant for the integration over paths. Furthermore, the new path variable $y(t)$ is restricted to take the value 0 at both end points. This substitution leads to a path integral independent of end-point positions.

At each t the variables x and y differ by the constant $\bar{x}$. (Of course, this is a different constant for each value of t.) Therefore, clearly, $dx_i = dy_i$ for each specific point t_i in the subdivision of time. In general, we may say $\mathfrak{D}x(t) = \mathfrak{D}y(t)$.

The integral for the action can be written

$$S[x(t)] = S[\bar{x}(t) + y(t)] = \int_{t_a}^{t_b} [a(t)(\dot{\bar{x}}^2 + 2\dot{\bar{x}}\dot{y} + \dot{y}^2) + \cdots] \, dt \qquad (3\text{-}48)$$

If all the terms which do not involve y are collected, the resulting integral is just $S[\bar{x}(t)] = S_{cl}$. If all the terms which contain y as a linear factor are collected, the resulting integral vanishes. This could be proved by actually carrying out the integration (some integration by parts would be involved); however, such a calculation is unnecessary, since we already know the result is true. The function $\bar{x}(t)$ is determined by this very requirement. That is, $\bar{x}$ is so chosen that there is no change in S, to first order, for variations of the path around $\bar{x}$. All that remain are the second-order terms in y. These can be easily picked out, so that we can write

$$S[x(t)] = S_{cl}[b,a] + \int_{t_a}^{t_b} [a(t)\dot{y}^2 + b(t)\dot{y}y + c(t)y^2] \, dt \qquad (3\text{-}49)$$

The integral over paths does not depend upon the classical path, so the kernel can be written

$$K(b,a) = e^{(i/\hbar)S_{cl}[b,a]} \int_0^0 \left(\exp\left\{ \frac{i}{\hbar} \int_{t_a}^{t_b} [a(t)\dot{y}^2 + b(t)\dot{y}y + c(t)y^2] \, dt \right\} \right) \mathfrak{D}y(t) \qquad (3\text{-}50)$$

Since all paths $y(t)$ start from and return to the point $y = 0$, the integral over paths can be a function only of times at the end points. This means that the kernel can be written as

$$K(b,a) = e^{(i/\hbar)S_{cl}[b,a]} F(t_a,t_b) \qquad (3\text{-}51)$$

so K is determined except for a function of t_a and t_b. In particular, its dependence upon the spatial variables x_a and x_b is completely worked out. It should be noted that the dependence of the kernel upon the coefficients of the linear terms $d(t)$ and $e(t)$ and the remaining coefficient $f(t)$ is also completely worked out.

This seems to be characteristic of various methods of doing path integrals; a great deal can be worked out by some general methods, but often a multiplying factor is not fully determined. It must be determined by some other known property of the solution, as, for example, by Eq. (2-31).

It is interesting to note that the approximate expression $K \sim e^{iS_{cl}/\hbar}$ is exact for the case that S is a quadratic form.

Problem 3-6 Since the free-particle lagrangian is a quadratic, show that (Prob. 2-1)

$$K(b,a) = F(t_b,t_a) \exp \frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)} \tag{3-52}$$

and give an argument to show that F can depend only on the difference $F(t_b - t_a)$.

Problem 3-7 Further information about F can be obtained from the property expressed by Eq. (2-31). First notice that the results of Prob. 3-6 imply that $F(t_b - t_a)$ can be written as $F(t)$, where t is the time interval $t_b - t_a$. By using this form for F in Eq. (3-52) and substituting into Eq. (2-31), express $F(t + s)$ in terms of $F(t)$ and $F(s)$, where $t = t_b - t_c$ and $s = t_c - t_a$. Show that if F is written as

$$F(t) = \sqrt{\frac{m}{2\pi i\hbar t}}\, f(t) \tag{3-53}$$

the new function $f(t)$ must satisfy

$$f(t + s) = f(t)f(s) \tag{3-54}$$

This means that $f(t)$ must be of the form

$$f(t) = e^{at} \tag{3-55}$$

where a may be complex, that is, $a = \alpha + i\beta$. It is difficult to obtain more information about the function $f(t)$ from the principles we have so far laid down. However, the special choice of the normalizing constant A defined in Eq. (2-21) implies that $f(\epsilon) = 1$ to first order in ϵ. This corresponds to setting a in Eq. (3-55) equal to 0. The resulting value of $F(t)$ is in agreement with Eq. (3-3).

It is clear from this example how the important properties of path integrals may be easily obtained even though the integrand may be a complicated function. So long as the integrand is an exponential function which contains the path variables only up to the second order, a solution that will be complete except possibly for some simple multiplying factors can be obtained. This is true regardless of the number of variables. Thus, for example, a path integral of the form

$$\int_a^b \int_c^d \cdots \int_k^l \exp\{E[x(t),y(t),\ \ldots\ ,z(t)]\} \mathfrak{D}x(t)\,\mathfrak{D}y(t)\ \cdots\ \mathfrak{D}z(t) \qquad (3\text{-}56)$$

contains as its important factor $e^{E_{cl}}$, where E_{cl} is the extremum of E subject to the boundary conditions. The only restriction is that in terms of the variables x, y, and so on, E is a function of the second degree. The remaining factor is a function of the times at the end points of the paths. For most of the path integrals which we shall study, the important information is contained in the exponential term rather than in the latter factor. In fact in most cases we shall not even find it necessary to evaluate this latter factor. This method of solving path integrals will be used frequently in the succeeding chapters.

3-6 MOTION IN A POTENTIAL FIELD

One simple application comes in the classical limiting case in which the action S is very large compared to Planck's constant $\hbar$. As we have already pointed out for this situation, the kernel K is approximately proportional to $e^{-iS_{cl}/\hbar}$. We can now see more mathematically the basis of this approximation. Only those paths quite near to the classical path $\bar{x}$ are important, so suppose we make the substitution $x = \bar{x} + y$. Now if the particle is moving through the potential $V(x)$, we can write

$$V(x) = V(\bar{x} + y) = V(\bar{x}) + yV'(\bar{x}) + \frac{y^2}{2}\,V''(\bar{x}) + \frac{y^3}{6}\,V'''(\bar{x}) + \cdots \qquad (3\text{-}57)$$

where the prime indicates differentiation with respect to x and all differentials are evaluated along the classical path $\bar{x}$. Only small values of y are important, so suppose V is a sufficiently smooth function that we can neglect terms of order y^3 and higher. This means that we assume that y^3V''' and all higher-order terms are negligible compared to the terms kept.

Under this assumption the integrand can be expressed as a quadratic form in y. In fact, since $\bar{x}$ makes S extreme, $S = S_{cl}$ + terms second order in y. The important term in the result is $e^{iS_{cl}/\hbar}$, where now, of course, S_{cl} contains the potential $V(\bar{x})$ along the classical path. The remaining integral on y goes from 0 to 0 and is of the form of the last factor in Eq. (3-50). It provides a smooth function as a factor to $e^{iS_{cl}/\hbar}$.

The result is true in situations other than the classical limiting case. For example, suppose V is a quadratic function of x. Then the solution is exact, since the expansion of V as in Eq. (3-57) contains no powers higher than the second. Some examples of this type are given in the problems. As another example, suppose V is a slowly varying function. In particular, if the third and higher derivatives are extremely small, the result given above is a very accurate approximation. This particular case is called the WKB approximation in quantum mechanics.

There are other situations in which the approximation is good. Suppose the total time interval for the motion is very short. If a particle moves along a path differing greatly from the classical path, it must have a very large extra velocity (to go out from the initial point and then return to the final point in the allotted time interval). The extra kinetic energy is proportional to the square of this large velocity, and the action contains a term roughly proportional to the kinetic energy multiplied by the time interval (thus, the square of the velocity multiplied by the time interval). The action for such paths will be very large, and the phase of the amplitude will vary greatly for closely neighboring paths. In this case again it is reasonable to drop the higher-order terms in the expansion of V.

Problem 3-8 For a harmonic oscillator the lagrangian is

$$L = \frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2 \tag{3-58}$$

Show that the resulting kernel is (see Prob. 2-2)

$$K = F(T)\exp\left\{\frac{im\omega}{2\hbar\sin\omega T}\left[(x_a^2 + x_b^2)\cos\omega T - 2x_ax_b\right]\right\} \tag{3-59}$$

where $T = t_b - t_a$. Note that the multiplicative function $F(T)$ has not been explicitly worked out. It can be obtained by other means, and for the harmonic oscillator it is (cf. Sec. 3-7)

$$F(T) = \left(\frac{m\omega}{2\pi i\hbar\sin\omega T}\right)^{1/2} \tag{3-60}$$

Problem 3-9 Find the kernel for a particle in a constant external field f where the lagrangian is

$$L = \frac{m}{2}\dot{x}^2 + fx \tag{3-61}$$

The result is

$$K = \left(\frac{m}{2\pi i\hbar T}\right)^{1/2} \exp\frac{i}{\hbar}\left[\frac{m(x_b - x_a)^2}{2T} + \frac{1}{2} fT(x_a + x_b) - \frac{fT^3}{24}\right] \tag{3-62}$$

where $T = t_b - t_a$.

Problem 3-10 The lagrangian for a particle of charge e and mass m in a constant external magnetic field B, in the z direction, is

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{eB}{2c}(x\dot{y} - y\dot{x}) \tag{3-63}$$

Show that the resulting kernel is

$$K = \left(\frac{m}{2\pi i\hbar T}\right)^{3/2}\left(\frac{\omega T/2}{\sin \omega T/2}\right)\exp\left(\frac{im\omega}{2\hbar}\left\{\frac{(z_b - z_a)^2}{T} + \frac{\omega}{2}\cot\frac{\omega T}{2}\,[(x_b - x_a)^2 + (y_b - y_a)^2] + \omega(x_a y_b - x_b y_a)\right\}\right) \tag{3-64}$$

where $T = t_b - t_a$, $\omega = eB/mc$.

Problem 3-11 Suppose the harmonic oscillator of Prob. 3-8 is driven by an external force $f(t)$. The lagrangian is

$$L = \frac{m}{2}\dot{x}^2 - \frac{m\omega^2}{2}x^2 + f(t)x \tag{3-65}$$

Show that the resulting kernel is

$$K = \sqrt{\frac{m\omega}{2\pi i\hbar \sin \omega T}}\, e^{(i/\hbar)S_{cl}}$$

where

$$S_{cl} = \frac{m\omega}{2\sin\omega T}\left[(\cos\omega T)(x_b^2 + x_a^2) - 2x_b x_a + \frac{2x_b}{m\omega}\int_{t_a}^{t_b} f(t)\sin\omega(t - t_a)\,dt + \frac{2x_a}{m\omega}\int_{t_a}^{t_b} f(t)\sin\omega(t_b - t)\,dt - \frac{2}{m^2\omega^2}\int_{t_a}^{t_b}\int_{t_a}^{t} f(t)f(s)\sin\omega(t_b - t)\sin\omega(s - t_a)\,ds\,dt\right] \tag{3-66}$$

with $T = t_b - t_a$.

This last result is of great importance in many advanced problems. It has particular applications in quantum electrodynamics because the electromagnetic field can be represented as a set of forced harmonic oscillators.

Problem 3-12 If the wave function for a harmonic oscillator is (at $t = 0$)

$$\psi(x,0) = \exp\left[-\frac{m\omega}{2\hbar}(x-a)^2\right] \tag{3-67}$$

then, using Eq. (3-42) and the results of Prob. 3-8, show that

$$\psi(x,T) = \exp\left\{-\frac{i\omega T}{2} - \frac{m\omega}{2\hbar}\left[x^2 - 2axe^{-i\omega T} + \tfrac{1}{2}a^2(1+e^{-2i\omega T})\right]\right\} \tag{3-68}$$

and find the probability distribution $|\psi|^2$.

3-7 SYSTEMS WITH MANY VARIABLES[1]

Suppose a system has several degrees of freedom. A kernel for such a system can be represented by the form of Eq. (2-25), where the symbol $x(t)$ now represents several coordinates rather than just one.

We take as a first example a particle moving in three dimensions. The path is defined by giving three functions $x(t)$, $y(t)$, and $z(t)$. The action for a free particle, for example, is

$$\frac{m}{2}\int_{t_a}^{t_b}[\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2]\,dt$$

The kernel to go from some initial point (x_a,y_a,z_a) at time t_a to a final point (x_b,y_b,z_b) at t_b is

$$K(x_b,y_b,z_b,t_b;x_a,y_a,z_a,t_a) = \int_a^b \left\{\exp\left[\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)\,dt\right]\right\} \mathcal{D}x(t)\,\mathcal{D}y(t)\,\mathcal{D}z(t) \tag{3-69}$$

The differential is written as $\mathcal{D}x(t)\,\mathcal{D}y(t)\,\mathcal{D}z(t)$. If the time is divided into intervals ϵ, the position at the time t_i is given by three variables x_i, y_i, z_i and the integral over all variables is dx_i, dy_i, dz_i for each i in an expression like Eq. (2-22). (More generally, if we represent the position by a vector **r** in some s-dimensional space, the differential at each point is the volume element dv_i or $d^s\mathbf{r}_i$ and we can write the product of these for each i as $\mathcal{D}^s\mathbf{r}_i$.)

---

[1] R. P. Feynman, Space-Time Approach to Non-relativistic Quantum Mechanics, *Rev. Mod. Phys.*, vol. 20, no. 2, p. 371, 1948.

If the definition of Eq. (2-22) is used, then the normalizing constant A [Eq. (2-21)] must be included for each variable in each time interval. Thus if the total time interval is broken up into N steps of length ϵ, the factor A^{-3N} must be included in the integral.

Another situation involving several variables is that of two interacting systems. Suppose one system consists of a particle of mass m with coordinates given by x and the other system contains a particle of mass M and coordinate X. Suppose these two systems interact through a potential $V(x,X)$. The resulting action is

$$S[x(t),X(t)] = \int_{t_a}^{t_b} \left[\frac{m}{2}\dot{x}^2 + \frac{M}{2}\dot{X}^2 - V(x,X)\right] dt \qquad (3\text{-}70)$$

so that the kernel is

$$K(x_b,X_b,t_b;x_a,X_a,t_a) = \int_a^b \int_a^b \left(\exp\left\{\frac{i}{\hbar} S[x(t),X(t)]\right\}\right) \mathcal{D}x(t)\,\mathcal{D}X(t) \qquad (3\text{-}71)$$

One might understand this generalization of Eq. (2-25) mathematically. Thus one might consider the motion of a point in some abstract two-dimensional space of coordinates x, X. However, it is much easier to think of it physically as representing the motion of two separate particles whose coordinates are respectively x and X. Then K is the amplitude that the particle of mass m goes from the point in space-time (x_a,t_a) to (x_b,t_b) and the particle of mass M goes from (X_a,t_a) to (X_b,t_b). The kernel is then the sum of an amplitude taken over all possible paths of both particles between their respective end points. The amplitude for any particular combination of paths [i.e., both $x(t)$ and $X(t)$ are specified] is $e^{iS/\hbar}$, where S is the action defined in Eq. (3-70). Mathematically, the amplitude is a functional of two independent functions $x(t)$ and $X(t)$, and the integral is over both of the variable functions.

3-8 SEPARABLE SYSTEMS

Suppose we have a situation in which two particles are present, both moving in one or perhaps more dimensions. Let the vector **x** represent the coordinates of one particle and the vector **X** represent the coordinates of the other, as in the paragraph above, except that now we extend the picture to a three-dimensional space. It may happen that the resulting action can be separated into two parts, as

$$S[\mathbf{x},\mathbf{X}] = S_x[\mathbf{x}] + S_X[\mathbf{X}] \tag{3-72}$$

where S_x involves only the paths $\mathbf{x}(t)$ and S_X involves only the paths $\mathbf{X}(t)$. This is the situation when the two particles do not interact.

In this case the kernel becomes the product of one factor depending on $\mathbf{x}$ and another depending on $\mathbf{X}$ alone. Thus

$$\begin{aligned} K(\mathbf{x}_b,\mathbf{X}_b,t_b;\mathbf{x}_a,\mathbf{X}_a,t_a) &= \int_a^b \int_a^b \left(\exp \frac{i}{\hbar}\{S_x[\mathbf{x}] + S_X[\mathbf{X}]\}\right) \mathcal{D}^3\mathbf{x}(t)\,\mathcal{D}^3\mathbf{X}(t) \\ &= \int_a^b \left(\exp\left\{\frac{i}{\hbar}S_x[\mathbf{x}]\right\}\right)\mathcal{D}\mathbf{x}(t) \int_a^b \left(\exp\left\{\frac{i}{\hbar}S_X[\mathbf{X}]\right\}\right)\mathcal{D}\mathbf{X}(t) \\ &= K_x(\mathbf{x}_b,t_b;\mathbf{x}_a,t_a)K_X(\mathbf{X}_b,t_b;\mathbf{X}_a,t_a) \end{aligned} \tag{3-73}$$

Here K_x is the amplitude computed as if only the particle of coordinates $\mathbf{x}$ were present, and K_X is defined similarly. Thus in a situation involving two independent noninteracting systems the kernel for an event involving both systems is the product of two independent kernels. These are the kernels for each particle to carry out its individual portion of the overall event.

The wave function in a situation involving several particles is defined in a straightforward manner by analogy with the corresponding kernel as $\psi(\mathbf{x},\mathbf{X}, \ldots, t)$. It is interpreted as the amplitude that, at time t, one particle is at the point $\mathbf{x}$, another particle is at the point $\mathbf{X}$, etc. The absolute square of the wave function is the probability that one particle is at the point $\mathbf{x}$ per unit volume, another particle is at the point $\mathbf{X}$ per unit volume, etc. Equation (3-42), which holds for the one-dimensional case, can be immediately extended to read

$$\begin{aligned} \psi(\mathbf{x},\mathbf{X}, \ldots, t) = \int\int K(\mathbf{x},\mathbf{X}, \ldots, t;\mathbf{x}',\mathbf{X}', \ldots, t') \\ \psi(\mathbf{x}',\mathbf{X}', \ldots, t')\, d\mathbf{x}'\, d\mathbf{X}' \end{aligned} \tag{3-74}$$

where $d\mathbf{x}'$ is the product of as many differentials as there are coordinates in $\mathbf{x}'$ space.

In case two independent particles are represented by the sets of coordinates $\mathbf{x}$ and $\mathbf{X}$, then the kernel K is the product of one function of $\mathbf{x}$ and t and another of $\mathbf{X}$ and t, as mentioned above. However, this does not imply that, in general, ψ is such a product. In the special case that ψ is at some particular time a product of a function of $\mathbf{x}$ and another of $\mathbf{X}$ [thus $\psi = f(\mathbf{x}) \cdot g(\mathbf{X})$], then it will remain so for all time. Each factor will change as it would for the partial system alone, since the kernel K represents the independent motion of two particles. But this is a special case. Just because the particles are independent now does not mean that they always were. There may have been some interaction in the past, which would imply that ψ is not a simple product.

Even though the action S does not appear as a simple product in the original coordinate system, there is often a transformation (such as that of center-of-gravity and internal coordinates) which will make it separable. Since the same form for the action is used in quantum mechanics as in classical physics, any transformation which will separate a classical system will also separate the corresponding quantum-mechanical system. Thus a part of the great body of work in classical physics can be applied directly to quantum mechanics. Such transformations are very important. It is hard to deal with a system consisting of several variables. Separation of variables permits us to reduce a complex problem to a number of simpler problems.

3-9 THE PATH INTEGRAL AS A FUNCTIONAL

When a problem contains more than one variable and a separation is not possible, the analysis is generally very difficult. Later on we shall discuss some approximations which can be applied in this case. Here we shall describe one very powerful tool which can sometimes be applied. Consider the kernel given by Eq. (3-71). This can be written out in detail as

$$K(b,a) = \int_a^b \int_a^b \left\{ \exp\left[\frac{i}{\hbar}\int_{t_a}^{t_b} \frac{m}{2}\dot{x}^2\,dt + \frac{i}{\hbar}\int_{t_a}^{t_b} \frac{M}{2}\dot{X}^2\,dt + \frac{i}{\hbar}\int_{t_a}^{t_b} V(x,X,t)\,dt \right]\right\} \mathcal{D}x(t)\,\mathcal{D}X(t) \qquad (3\text{-}75)$$

First, suppose we carry out the integral over the paths $X(t)$. The result can be written formally as

$$K(b,a) = \int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b} \frac{m}{2}\dot{x}^2\,dt\right)\right] T[x(t)]\,\mathcal{D}x(t) \qquad (3\text{-}76)$$

where

$$T[x(t)] = \int_a^b \left(\exp\left\{\frac{i}{\hbar}\int_{t_a}^{t_b}\left[\frac{M}{2}\dot{X}^2 + V(x,X,t)\right]\right\}\right) dt\,\mathcal{D}X(t) \qquad (3\text{-}77)$$

These results are interpreted in the following manner. Integrating over all paths available to the X particle produces a *functional* T. A functional is a number whose value depends on specifying a complete function. For example, the area under a curve is a functional of the curve $A = \int f(y)\,dy$. To find it, a function (the curve) must be specified. We write a functional as $A[f(y)]$ to indicate that A depends on the function $f(y)$. We do not write $A(f(y))$, for that might be interpreted as a function of a function, i.e., that A just depends on what value f takes at some specified point y. This is not the case. $A[f(y)]$ depends on the entire shape of the function $f(y)$. It does not depend on y in any way.

The functional defined in Eq. (3-77) is the amplitude that the X particle alone goes between its end points X_a and X_b under the influence of a potential V. This potential, which depends upon both x and X, is computed assuming x is held to be a fixed path as X changes. Thus it is the potential for the X particle when the x particle is moving along a specific trajectory. Clearly, this amplitude T depends upon the trajectory chosen for $x(t)$, so we write it as a functional of $x(t)$. Then the total amplitude is obtained by summing over all paths a functional consisting of the product of T and the free-particle kernel for $x(t)$.

Thus the amplitude K, like all others, is a sum over the amplitudes of all possible alternatives. Each of these amplitudes is a product of two lesser amplitudes. The first of these is the amplitude T that the X particle goes between its given end points when x has a specified trajectory. The second is the amplitude that x has that specified trajectory. The final sum over alternatives becomes the sum over all possible trajectories of x. It is important to understand this concept clearly, for it includes one of the fundamental principles of quantum electrodynamics, a subject which will be taken up in a later chapter.

Of course it is not practical to use this method unless the integral T can actually be worked out, either exactly or approximately, for the possible values of the trajectory $x(t)$. As we have seen (cf. Prob. 3-11) one exact case is that in which X is a harmonic oscillator. This is a very important practical case. For example, when a particle interacts with a quantized field, the field is an oscillator.

3-10 INTERACTION OF A PARTICLE AND A HARMONIC OSCILLATOR

We shall consider in more detail the interaction of a particle and a harmonic oscillator. Let the coordinates of the particle be $\mathbf{x}$ and those of the oscillator be $\mathbf{X}$. The action can be written as

$$S[\mathbf{x},\mathbf{X}] = S_0[\mathbf{x}] + \int_{t_a}^{t_b} g[\mathbf{x}(t),t]\mathbf{X}(t)\,dt + \int_{t_a}^{t_b} \frac{M}{2}(\dot{\mathbf{X}}^2 + \omega^2\mathbf{X}^2)\,dt \tag{3-78}$$

where S_0 is the action of the particle in the absence of the oscillator. In the discussion above we assumed that this action corresponded to that for a free particle. This assumption is not necessary. The motion of $\mathbf{x}$ could be complicated by the existence of a potential depending upon $\mathbf{x}$ and t only. Thus, for example, the action S_0 might be

$$S_0[\mathbf{x}] = \int_{t_a}^{t_b} \left[\frac{m}{2}\dot{\mathbf{x}}^2 - V(\mathbf{x},t)\right] dt \tag{3-79}$$

The second term in Eq. (3-78) represents the interaction between the particle and the oscillator. Note that this term is linear in $\mathbf{X}$. Omission of a dependence upon $\dot{\mathbf{X}}$ does not imply any loss in generality, since if such a term were to occur, it could be removed by an integration by parts. We can call the coefficient g the coupling coefficient. Its dependence upon $\mathbf{x}(t)$ is indicated, but it could also depend upon other variables, such as $\dot{\mathbf{x}}(t)$. Since the analysis we are presenting is general, it is not important to write down the exact form of g. The last term in Eq. (3-78) is, of course, the action of the oscillator alone. By combining this with the second term, the function T of Eq. (3-77) can be written as

$$T[\mathbf{x}(t)] = \int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\left\{\frac{M}{2}(\dot{\mathbf{X}}^2 - \omega^2\mathbf{X}^2) + g[\mathbf{x}(t),t]\mathbf{X}(t)\right\} dt\right)\right] \mathcal{D}\mathbf{X}(t) \tag{3-80}$$

Now as far as $\mathbf{X}$ is concerned, the situation is just that of a forced harmonic oscillator. The forcing function $g[\mathbf{x}(t),t]$ is some special function of t, say, $f(t)$. Thus the path integral is the same as that considered in Prob. 3-11, with $f(t)$ replaced by $g[\mathbf{x}(t),t]$ and the initial and final coordinate values (x_b,x_a) replaced by $(\mathbf{X}_b,\mathbf{X}_a)$.

For illustrative purposes, to simplify the expressions somewhat, we take the special case in which the oscillator initially and finally is at the origin, so $\mathbf{X}_b = \mathbf{X}_a = 0$ (the general case is just as easily handled). Then according to Prob. 3-11 in this case we have

$$T = \left(\frac{m\omega}{2\pi\hbar i \sin\omega T}\right)^{1/2} \exp\left\{\frac{i}{\hbar m\omega \sin\omega T}\int_{t_a}^{t_b}\int_{t_a}^{t} g[\mathbf{x}(t),t]g[\mathbf{x}(s),s] \sin\omega(t_b - t)\sin\omega(s - t_a)\,ds\,dt\right\} \tag{3-81}$$

Therefore, the kernel for the present situation can be written

$$K(b,a) = \left(\frac{m\omega}{2\pi\hbar i \sin \omega T}\right)^{1/2} \int_a^b \left(\exp \frac{i}{\hbar} \left\{\frac{m}{2} \int_{t_a}^{t_b} \dot{\mathbf{x}}(t)^2 \, dt - \frac{1}{m\omega \sin \omega T} \int_{t_a}^{t_b} \int_{t_a}^{t} g[\mathbf{x}(t),t] g[\mathbf{x}(s),s] \sin \omega(t_a - t) \sin \omega(s - t_a) \, ds \, dt\right\}\right) \mathcal{D}\mathbf{x}(t) \quad (3\text{-}82)$$

with a similar (but more complicated) expression for arbitrary $\mathbf{X}_a$, $\mathbf{X}_b$.

This is a more complicated path integral than any we have had to solve so far. It is not possible to proceed further with the solution until various methods of approximation have been developed in succeeding chapters. Note that the integrand of this path integral can still be thought of as being of the form $e^{(i/\hbar)S}$, but now S is no longer a function of only $\dot{\mathbf{x}}$, $\mathbf{x}$, and t. Instead, S contains a product of variables defined at two different times, s and t. The separation of past and future can no longer be made. This happens because the variable $\mathbf{x}$ at some previous time affects the oscillator which, at some later time, reacts back to affect $\mathbf{x}$. No wave function $\psi(\mathbf{x},t)$ can be defined to give the amplitude that the particle is at some particular place $\mathbf{x}$ at a particular time t. Such an amplitude would be insufficient for continuing calculations into the future, since at any time one must also know what the oscillator is doing.

3-11 EVALUATION OF PATH INTEGRALS BY FOURIER SERIES

Consider the path integral for the harmonic oscillator problem (Prob. 3-8). This is

$$K(b,a) = \int_a^b \left(\exp \left\{\frac{i}{\hbar} \int_{t_a}^{t_b} \left[\frac{m}{2} (\dot{x}^2 - \omega^2 x^2) \, dt\right]\right\}\right) \mathcal{D}x(t) \quad (3\text{-}83)$$

Using the methods of Sec. 3-5 this path integral can be reduced to a product of two functions, as in Prob. 3-8. The more important of these two functions depends upon the classical orbit for a harmonic oscillator and is given in Eq. (3-59). The remaining function depends upon the time interval only and is written down in Eq. (3-60). This latter function can be written as

$$F(T) = \int_0^0 \left\{\exp \left[\frac{i}{\hbar} \int_0^T \frac{m}{2} (\dot{y}^2 - \omega^2 y^2) \, dt\right]\right\} \mathcal{D}y(t) \quad (3\text{-}84)$$

We shall solve this, at least to within a factor independent of ω, by a method which illustrates still another way of handling path integrals. Since all paths $y(t)$ go from 0 at $t = 0$ to 0 at $t = T$, such paths can be written as a Fourier sine series with a fundamental period of T. Thus

$$y(t) = \sum_n a_n \sin \frac{n\pi t}{T} \tag{3-85}$$

It is possible then to consider the paths as functions of the coefficients of a_n instead of functions of y at any particular value of t. This is a linear transformation whose jacobian J is a constant, obviously independent of ω, m, and $\hbar$.

Of course, it is possible to evaluate this jacobian directly. However, here we shall avoid the evaluation of J by collecting all factors which are independent of ω (including J) into a single constant factor. We can always recover the correct factor at the end, since we know the value for $\omega = 0$, $F(T) = \sqrt{m/2\pi i\hbar T}$ (a free particle).

The integral for the action can be written in terms of the Fourier series of Eq. (3-85). Thus the kinetic-energy term becomes

$$\begin{aligned}\int_0^T \dot{y}^2\,dt &= \sum_n \sum_m \frac{n\pi}{T}\frac{m\pi}{T} a_n a_m \int_0^T \cos\frac{n\pi t}{T}\cos\frac{m\pi t}{T}\,dt \\ &= T\cdot\frac{1}{2}\sum_n \left(\frac{n\pi}{T}\right)^2 a_n^2 \end{aligned}\tag{3-86}$$

and similarly the potential-energy term is

$$\int_0^T y^2\,dt = T\cdot\frac{1}{2}\sum_n a_n^2 \tag{3-87}$$

On the assumption that the time T is divided into discrete steps of length ϵ as for Eq. (2-19) so that there are only a finite number N of coefficients a_n, the path integral becomes

$$F(T) = J\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\exp\left\{\sum_{n=1}^{N}\frac{im}{2\hbar}\left[\left(\frac{n\pi}{T}\right)^2 - \omega^2\right]a_n^2\right\}\right)\frac{da_1}{A}\frac{da_2}{A}\cdots\frac{da_N}{A} \tag{3-88}$$

Since the exponent can be separated into factors, the integral over each coefficient a_n can be done separately. The result of one such integration is

$$\int_{-\infty}^{\infty} \left\{ \exp \left[\frac{im}{2\hbar} \left(\frac{n^2\pi^2}{T^2} - \omega^2 \right) a_n{}^2 \right] \right\} \frac{da_n}{A} = \left(\frac{n^2\pi^2}{T^2} - \omega^2 \right)^{-\frac{1}{2}} \tag{3-89}$$

Thus the path integral is proportional to

$$\prod_{n=1}^{N} \left(\frac{n^2\pi^2}{T^2} - \omega^2 \right)^{-\frac{1}{2}} = \prod_{n=1}^{N} \left(\frac{n^2\pi^2}{T^2} \right)^{-\frac{1}{2}} \prod_{n=1}^{N} \left(1 - \frac{\omega^2 T^2}{n^2\pi^2} \right)^{-\frac{1}{2}} \tag{3-90}$$

The first product does not depend on ω and combines with the jacobian and other factors we have collected into a single constant. The second factor has the limit $[(\sin \omega T)/\omega T]^{-\frac{1}{2}}$ as $N \to \infty$, that is, as $\epsilon \to 0$. Thus

$$F(T) = C \left(\frac{\sin \omega T}{\omega T} \right)^{-\frac{1}{2}} \tag{3-91}$$

where C is independent of ω. But for $\omega = 0$ our integral is that for a free particle, for which we have already found that

$$F(T) = \left(\frac{m}{2\pi i \hbar T} \right)^{\frac{1}{2}} \tag{3-92}$$

Hence for the harmonic oscillator we have

$$F(T) = \left(\frac{m\omega}{2\pi i \hbar \sin \omega T} \right)^{\frac{1}{2}} \tag{3-93}$$

which is to be substituted in Eq. (3-59) to obtain the complete solution.

Problem 3-13 By keeping track of all the constants, show that the implication is that the jacobian satisfies

$$J \sqrt{N} \left(\frac{T}{\pi} \right)^N \prod_{n=1}^{N} \frac{1}{n} \to 1 \qquad \text{as } N \to \infty \tag{3-94}$$

4

The Schrödinger Description of Quantum Mechanics

THE path integrals which we have discussed so far [with the exception of Eq. (3-82)] have integrands which are exponentials of actions with the property

$$S[2,1] = S[2,3] + S[3,1] \tag{4-1}$$

Such path integrals can be analyzed in terms of the properties of integral equations which can be deduced from them. We have already seen this in Chap. 2 [e.g., Eq. (2-31)] and Chap. 3 [e.g., Eq. (3-42)].

A still more convenient method is to reduce the path integrals to differential equations if possible. This possibility exists in quantum mechanics and is, in fact, the most convenient way to present that theory. It is in almost every case easier to solve the differential equation than it is to evaluate the path integral directly. The conventional presentation of quantum mechanics is based on this differential equation, called the Schrödinger equation. Here we shall derive this equation from our formulation. We shall not solve this equation for a large number of examples, because such solutions are presented in a detailed and satisfactory fashion in other books on quantum mechanics.[1]

In this chapter our purpose is twofold: (1) For the reader primarily interested in quantum mechanics our aim is to connect the path integral formulation with other formulations which are found in the standard literature and textbooks so that he can continue his study in those books and can learn to translate back and forth between the two different languages. (2) For the reader primarily interested in path integrals this chapter will show a technique which is available for a certain class of path integrals to reduce these path integrals to differential equations. This technique is best shown by the particular example of quantum mechanics which we shall develop here.

4-1 THE SCHRÖDINGER EQUATION

The Differential Equation Form. The reason that we can develop a differential equation is that the relationship of Eq. (4-1) is correct for any values of the points 1, 2, and 3. For example, the time t_2 can be only an infinitesimal time ϵ greater than the time t_3. This will permit us to relate the value of a path integral for one time to its value a short time later. In this manner we can obtain a differential equation for the path integral.

---

[1] For example, see L. I. Schiff, "Quantum Mechanics," 2d ed., McGraw-Hill Book Company, New York, 1955.

We have already found that as a consequence of Eq. (4-1) we can define a wave function. Furthermore, we know that the equation

$$\psi(x_2,t_2) = \int_{-\infty}^{\infty} K(x_2,t_2;x_1,t_1)\psi(x_1,t_1)\,dx_1 \tag{4-2}$$

gives the wave function at a time t_2 in terms of the wave function at a time t_1. In order to obtain the differential equation that we seek, we apply this relationship in the special case that the time t_2 differs only by an infinitesimal interval ϵ from t_1. The kernel $K(2,1)$ is proportional to the exponential of $i/\hbar$ times the action for the interval t_1 to t_2. For a short interval ϵ the action is approximately ϵ times the lagrangian for this interval. That is, using the same approximation as that of Eq. (2-34), we have

$$\psi(x, t + \epsilon) = \int_{-\infty}^{\infty} \frac{1}{A} \exp\left[\epsilon \frac{i}{\hbar} L\left(\frac{x-y}{\epsilon}, \frac{x+y}{2}\right)\right] \psi(y,t)\,dy \tag{4-3}$$

We shall now apply this to the special case of a particle moving in a potential $V(x,t)$ in one dimension, i.e., that for which $L = m\dot{x}^2/2 - V(x,t)$. In this case Eq. (4-3) becomes

$$\psi(x, t + \epsilon) = \int_{-\infty}^{\infty} \frac{1}{A} \left\{\exp\left[\frac{i}{\hbar}\frac{m(x-y)^2}{\epsilon}\right]\right\} \left\{\exp\left[-\frac{i}{\hbar}\epsilon V\left(\frac{x+y}{2}, \epsilon t\right)\right]\right\} \psi(y,t)\,dy \tag{4-4}$$

The quantity $(x - y)^2/\epsilon$ appears in the exponent of the first factor. It is clear that if y is appreciably different from x, this quantity is very large and the exponential consequently oscillates very rapidly as y varies. When this factor oscillates rapidly, the integral over y gives a very small value (because of the smooth behavior of the other factors). Only if y is near x (where the exponential changes more slowly) do we get important contributions. For this reason we make the substitution $y = x + \eta$ with the expectation that appreciable contributions to the integral will occur only for small η. We obtain

$$\psi(x, t + \epsilon) = \int_{-\infty}^{\infty} \frac{1}{A} e^{im\eta^2/2\hbar\epsilon} e^{-(i\epsilon/\hbar)V[(x+\eta)/2,t]} \psi(x + \eta, t)\,d\eta \tag{4-5}$$

The phase of the first exponential changes by the order of 1 radian when η is of the order $\sqrt{\epsilon\hbar/m}$, so that most of the integral is contributed by values of η in this order.

We may expand ψ in a power series. We need only keep terms of order ϵ. This implies keeping second-order terms in η. The term

$\epsilon V[(x+\eta)/2,\, t]$ may be replaced by $\epsilon V(x,t)$ because the error is of higher order than ϵ. Expanding the left-hand side to first order in ϵ and the right-hand side to first order in ϵ and second order in η, we obtain

$$\psi(x,t) + \epsilon \frac{\partial \psi}{\partial t} = \int_{-\infty}^{\infty} \frac{1}{A} e^{im\eta^2/2\hbar\epsilon} \left[1 - \frac{i\epsilon}{\hbar} V(x,t) \right] \times \left[\psi(x,t) + \eta \frac{\partial \psi}{\partial x} + \frac{1}{2} \eta^2 \frac{\partial^2 \psi}{\partial x^2} \right] d\eta \tag{4-6}$$

If we take the leading term on the right-hand side, we have the quantity $\psi(x,t)$ multiplied by the integral

$$\frac{1}{A} \int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\epsilon}\, d\eta = \frac{1}{A} \left(\frac{2\pi i\hbar\epsilon}{m} \right)^{1/2} \tag{4-7}$$

On the left-hand side we have just $\psi(x,t)$. In order that both sides agree in the limit ϵ approaches 0, it is necessary that A be so chosen that the expression of Eq. (4-7) equals 1. That is,

$$A = \left(\frac{2\pi i\hbar\epsilon}{m} \right)^{1/2} \tag{4-8}$$

as we have seen previously [cf. Eq. (2-21)]. This is a way of obtaining the quantity A in more complicated problems also. The A must be so chosen that the equation is correct to zero order in ϵ. Otherwise, no limit will exist as ϵ approaches 0 in the original path integral.

In order to evaluate the right-hand side of Eq. (4-6), we shall have to use two integrals

$$\int_{-\infty}^{\infty} \frac{1}{A} e^{im\eta^2/2\hbar\epsilon} \eta\, d\eta = 0 \tag{4-9}$$

and

$$\int_{-\infty}^{\infty} \frac{1}{A} e^{im\eta^2/2\hbar\epsilon} \eta^2\, d\eta = \frac{i\hbar\epsilon}{m} \tag{4-10}$$

Writing out the right-hand side of Eq. (4-6) gives

$$\psi + \epsilon \frac{\partial \psi}{\partial t} = \psi - \frac{i\epsilon}{\hbar} V\psi - \frac{\hbar\epsilon}{2im} \frac{\partial^2 \psi}{\partial x^2} \tag{4-11}$$

This will be true to order ϵ if ψ satisfies the differential equation

$$-\frac{\hbar}{i} \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x,t)\psi \tag{4-12}$$

This is the Schrödinger equation for our problem of a particle moving in one dimension. Corresponding equations in more complicated situations can be worked out in the same way, as demonstrated by the following problems.

Problem 4-1 Show that for a single particle moving in three dimensions in a potential V the Schrödinger equation is

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \tag{4-13}$$

This equation was discovered by Schrödinger in 1925 and formed the central feature of the development of quantum mechanics thereafter.

The Operator Form. The equations which result from various problems corresponding to different forms for the lagrangian can all be written for convenience in the form

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = H\psi \tag{4-14}$$

Here H does not represent a number but indicates an operation on ψ. It is called the hamiltonian operator. For example, in Eq. (4-12) this operation is

$$H = -\frac{\hbar^2}{2m^2}\frac{\partial^2}{\partial x^2} + V \tag{4-15}$$

Such an equation with operators on both sides means this: If any function f is written after each operator on each side, the equation will be true. That is, Eq. (4-15) symbolizes the statement: The relation

$$Hf = -\frac{\hbar^2}{2m^2}\frac{\partial^2 f}{\partial x^2} + Vf \tag{4-16}$$

holds for any function f.

Problem 4-2 Show that for a charged particle in a magnetic field for which the lagrangian is

$$L = \frac{m\dot{r}^2}{2} + \frac{e}{c}\dot{\mathbf{r}}\cdot\mathbf{A} - e\phi \tag{4-17}$$

where $\dot{\mathbf{r}}$ is the velocity vector, e is the charge, c is the velocity of light, and $\mathbf{A}$ and ϕ are the vector and scalar potentials, the corresponding

Schrödinger equation is

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\cdot\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\psi + e\phi\psi \tag{4-18}$$

Thus the hamiltonian is

$$H = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right)\cdot\left(\frac{\hbar}{i}\nabla - \frac{e}{c}\mathbf{A}\right) + e\phi \tag{4-19}$$

Problem 4-3 Show that the complex conjugate function ψ^*, defined as the function ψ with every i changed to $-i$, satisfies

$$\frac{\hbar}{i}\frac{\partial\psi^*}{\partial t} = (H\psi)^* \tag{4-20}$$

The notation for operators can be described by giving a number of examples. For example, the operator x means multiplication by x, the operator x^2 means multiplication by x^2, the operator $V(x)$ (some function of x) means multiplication by $V(x)$, the operator $\partial/\partial x$ means partial differentiation with respect to x, $\partial\psi/\partial x$, etc.

If A and B are operators, then the operator AB means that we first apply B and then A, that is, $AB\psi$ means $A(B\psi)$. Thus, for example, the operator $x(\partial/\partial x)$ means x times $\partial\psi/\partial x$. On the other hand, the operator $(\partial/\partial x)x$ means the partial derivative with respect to x of $x\psi$, or $(\partial/\partial x)(x\psi) = x(\partial\psi/\partial x) + \psi$.

We see that in general the operator AB and the operator BA are *not* identical. We further define the operator $A + B$ by the rule that $A + B$ operating on ψ is $A\psi + B\psi$. For example, the previous equation can be written as an equation among operators as follows:

$$\frac{\partial}{\partial x}x = x\frac{\partial}{\partial x} + 1 \tag{4-21}$$

the meaning being that $(\partial/\partial x)xf = x(\partial/\partial x)f + f$ for any function f.

Problem 4-4 Show

$$\frac{\partial^2}{\partial x^2}x = x\frac{\partial^2}{\partial x^2} + 2\frac{\partial}{\partial x} \tag{4-22}$$

and therefore that, for the H of Eq. (4-15),

$$Hx - xH = -\frac{\hbar^2}{m}\frac{\partial}{\partial x} \tag{4-23}$$

This operator notation is used a great deal in the conventional formulations of quantum mechanics.

The Schrödinger Equation for the Kernel. Since $K(2,1)$, thought of as a function of the variables 2, is a special wave function (namely, that for a particle which starts at 1), we see that K must also satisfy a Schrödinger equation. Thus for the case specified by Eq. (4-15)

$$-\frac{\hbar}{i}\frac{\partial}{\partial t_2}K(2,1) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_2{}^2}K(2,1) + V(2)K(2,1) \tag{4-24}$$

for $t_2 > t_1$. In general we have

$$-\frac{\hbar}{i}\frac{\partial K(2,1)}{\partial t_2} - H_2K(2,1) = 0 \qquad \text{for } t_2 > t_1 \tag{4-25}$$

wherein the operator H_2 operates on the variables 2 only.

Problem 4-5 Using the relation

$$K(2,1) = \int_{-\infty}^{\infty} K(2,3)K(3,1)\,dx_3 \tag{4-26}$$

with $t_3 - t_1 = \epsilon$, an infinitesimal, show that if t_2 is greater than t_1, the kernel K satisfies

$$+\frac{\hbar}{i}\frac{\partial}{\partial t_1}K(2,1) - H_1^*K(2,1) = 0 \tag{4-27}$$

where H_1 now operates on variables 1 only.

The function $K(2,1)$ defined by a path integral in Eq. (2-25) is defined only for $t_2 > t_1$. The function is not defined if $t_2 < t_1$. It will prove to be very convenient in later work (e.g., Chap. 6) to define $K(2,1)$ to be zero for $t_2 < t_1$. [With this convention Eq. (4-2), for example, is valid only if $t_2 > t_1$.] With the condition

$$K(2,1) = 0 \qquad \text{for } t_2 < t_1 \tag{4-28}$$

it is evident that Eq. (4-25) is satisfied also for $t_2 < t_1$ (in a trivial fashion, of course, since $K = 0$). But this equation is not satisfied at the point $t_2 = t_1$ because $K(2,1)$ is discontinuous at $t_2 = t_1$.

Problem 4-6 Show that $K(2,1) \to \delta(x_2 - x_1)$ as $t_2 \to t_1 + 0$.

From the result of Prob. 4-6 we see that the derivative of K with respect to t_2 gives a delta function in the time multiplied by the height of the jump, $\delta(x_2 - x_1)$. Hence $K(2,1)$ satisfies

$$-\frac{\hbar}{i}\frac{\partial K(2,1)}{\partial t_2} - H_2K(2,1) = -\frac{\hbar}{i}\,\delta(x_2 - x_1)\delta(t_2 - t_1) \tag{4-29}$$

This equation plus the boundary condition of Eq. (4-28) could serve to define $K(2,1)$ if one were to have started out from the Schrödinger equation as the fundamental definition in quantum mechanics. It is clear that the quantity $K(2,1)$ is a kind of Green's function for the Schrödinger equation.

The Conservation of Probability. The hamiltonian operator given by Eq. (4-15) has the interesting property that, if f, g are any functions which fall off to zero at infinity,

$$\int_{-\infty}^{\infty} (Hg)^*f\,dx = \int_{-\infty}^{\infty} g^*(Hf)\,dx \tag{4-30}$$

The meaning of the symbols is this. On the left we are to take g, operate on it with H (forming Hg), and then take the complex conjugate. The result is then multiplied by f and integrated over all space. The result is the same as taking Hf, multiplying by the complex conjugate of g, and integrating. It is easily verified that this is true by integrating the term $\int (Hg)^*f\,dx$ (by parts, where necessary).

For our example in Eq. (4-15) we have for the left side of Eq. (4-30)

$$\begin{aligned} -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty}\frac{d^2g^*}{dx^2}f\,dx + \int_{-\infty}^{\infty} Vg^*f\,dx &= -\frac{\hbar^2}{2m}\left(\frac{dg^*}{dx}f - g^*\frac{df}{dx}\right)\Big|_{-\infty}^{+\infty} \\ &\quad -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} g^*\frac{d^2f}{dx^2}\,dx + \int_{-\infty}^{\infty} Vg^*f\,dx \end{aligned} \tag{4-31}$$

(integrating by parts twice). If f, g fall off at infinity, the integrated parts vanish and Eq. (4-30) is established. An operator which has the property given by Eq. (4-30) is called *hermitian*. In all cases of quantum mechanics the hamiltonian is hermitian. For more general cases than that considered above the integration over our one-dimensional variable x becomes an integration (or sum) over all the variables of the system.

If we put f and g equal to $\psi(x,t)$, we get

$$\int (H\psi)^*\psi\,dx = \int \psi^*(H\psi)\,dx \tag{4-32}$$

and if ψ satisfies the wave equation (4-14), this becomes

$$\int \frac{\partial\psi^*}{\partial t}\psi\,dx + \int \psi^*\frac{\partial\psi}{\partial t}\,dx = \frac{d}{dt}\left(\int \psi^*\psi\,dx\right) = 0 \tag{4-33}$$

That is, $\int\psi^*\psi\,dx$ is a constant independent of time. This is easily interpreted. For if ψ is suitably normalized, $\psi^*\psi$ is the probability of being found at x; so the integral is the probability of being found somewhere, which is certainty (or 1) and is constant. Of course, as far as the wave equation is concerned ψ can be multiplied by any constant and still be a solution. Then $\psi^*\psi$ is multiplied by the square of the constant, and the integral is this constant squared.

It is fundamental to our definition of ψ as probability amplitude that the integral of $\psi^*\psi$ is constant. In terms of the kernel this means that if f is the wave function at time t_1, then at time t_2 it has the same square integral. That is, if

$$\psi(2) = \int K(2,1)f(1)\,dx_1 \tag{4-34}$$

$$\int \psi^*(2)\psi(2)\,dx_2 = \int f^*(1)f(1)\,dx_1 \tag{4-35}$$

or

$$\int\int\int K^*(2;x_1',t_1)K(2;x_1,t_1)f^*(x_1')f(x_1)\,dx_1\,dx_1'\,dx_2 = \int f^*(x_1)f(x_1)\,dx_1 \tag{4-36}$$

For this to be true for arbitrary f we must have

$$\int K^*(2;x_1',t_1)K(2;x_1,t_1)\,dx_2 = \delta(x_1' - x_1) \tag{4-37}$$

That is, in order to interpret ψ as a probability amplitude, the kernel must satisfy Eq. (4-37). We have derived this by means of the Schrödinger equation. It would be nicer to demonstrate this and other properties, such as Eq. (4-38) and Prob. 4-7, directly in terms of the path integral definition of K instead of coming through the differential equation. It is possible, of course, but it is not so simple or neat as a derivation of such a fundamental relation should be. One can verify Eq. (4-37) as follows: For a small interval with $t_1 = t_2 - \epsilon$, Eq. (4-37) follows directly from the expression $e^{i\epsilon L/\hbar}$ for this interval. By induction, the complete Eq. (4-37) results. One disadvantage ot the approach to quantum mechanics through the path integrals is the fact that relations involving ψ^* or K^* are not self-evident.

By multiplying Eq. (4-37) by $K(1,3)$ and integrating on x_1, we can show

$$\int K^*(2,1)K(2,3)\,dx_2 = K(1,3) \tag{4-38}$$

where $t_2 > t_1 > t_3$. Compare this to $\int K(1,2)K(2,3)\,dx_2 = K(1,3)$ for $t_1 > t_2 > t_3$. We may describe the second relation this way: Starting at t_3, $K(2,3)$ gives us the amplitude at the later time t_2. If we wish to go to a still later time t_1, we can do so by using the kernel $K(1,2)$. On the other hand, if having the amplitude at t_2 we want to *work back* to find it at an earlier time $t_1 < t_2$, we can do this by using the function $K^*(2,1)$ according to Eq. (4-38). That is, we can say $K^*(2,1)$ undoes the work of $K(1,2)$.

Problem 4-7 Show that if $t_1 < t_3$, the left-hand side of Eq. (4-38) equals $K^*(3,1)$.

4-2 THE TIME - INDEPENDENT HAMILTONIAN

Steady States of Definite Energy. The special case that the hamiltonian H is independent of time is of great practical importance. This corresponds to the case that the action S does not depend on the time explicitly; i.e., the potentials **A** and V, for example, do not contain t. In this case the kernel cannot depend upon the absolute time but instead is a function only of the interval $t_2 - t_1$. As a consequence of this fact, there exist wave functions that depend periodically on the time.

It is easiest to see what happens by studying the differential equation. Starting from the Schrödinger equation (4-14), we try a special solution of the form $\psi = f(t)\phi(x)$, a function of a time only multiplied by a function of position only. Substitution gives us the relation

$$-\frac{\hbar}{i}f'(t)\phi(x) = Hf(t)\phi(x) = f(t)H\phi(x) \tag{4-39}$$

or

$$-\frac{\hbar}{i}\frac{f'}{f} = \frac{1}{\phi}H\phi \tag{4-40}$$

The left-hand side of this equation does not depend upon x, whereas the right-hand side is independent of t. If they are always equal, neither side can depend upon either variable t or x. That is, each side must be a constant. Let us call this constant E. Then $f' = -(i/\hbar)Ef$, or $f = e^{-iEt/\hbar}$ with an arbitrary constant factor. Thus the special solution is of the form

$$\psi(x,t) = e^{-(i/\hbar)Et}\phi(x) \tag{4-41}$$

where ϕ satisfies

$$H\phi = E\phi \tag{4-42}$$

That is, for this special solution the wave function oscillates with a definite frequency. We saw that the frequency with which a wave function oscillates corresponds, in classical physics, to the energy.

Therefore, we say that when the wave function is of this special form, the state has a definite energy E. For each value of E a different particular function ϕ [a solution of Eq. (4-42)] must be sought.

The probability that a particle is at x is the absolute square of the wave function ψ, or $|\psi|^2$. In view of Eq. (4-41) this is equal to $|\phi|^2$ and does not depend upon the time. That is, the probability of finding the particle in any location is independent of the time. We say under these circumstances that the system is in a stationary state—stationary in the sense that there is no variation in the probabilities as a function of time.

This situation is somewhat related to the uncertainty principle; for in a situation in which we know that the energy is exactly E we must be completely uncertain of the time. This is consonant with the idea that the properties of an atom in a specific state are absolutely independent of the time, so that at any time we would obtain the same result.

Suppose that E_1 is a possible energy for which Eq. (4-42) has a solution ϕ_1 and that E_2 is another value for energy for which this equation has some other solution ϕ_2. Then we know two special solutions of the Schrödinger equation, namely,

$$\psi_1 = e^{-(i/\hbar)E_1t}\phi_1(x) \qquad \text{and} \qquad \psi_2 = e^{-(i/\hbar)E_2t}\phi_2(x) \tag{4-43}$$

Since the Schrödinger equation is linear, it is clear that if ψ is a solution, so is $c\psi$. Furthermore, if ψ_1 is a solution and ψ_2 is a solution, then the sum of ψ_1 and ψ_2 is also a solution. Evidently, then, the function

$$\psi = c_1e^{-(i/\hbar)E_1t}\phi_1(x) + c_2e^{-(i/\hbar)E_2t}\phi_2(x) \tag{4-44}$$

is also a solution of the Schrödinger equation.

As a matter of fact, it can be shown that if all of the possible values of E and the corresponding functions ϕ are worked out, any solution of ψ of Eq. (4-14) can be written as a linear combination of these special solutions of definite energy.

The total probability to be anywhere is constant, as we showed in Sec. 4-1. This must be true no matter what the values of c_1 and c_2, so that, using Eq. (4-44) for ψ, we have

$$\begin{aligned}\int\psi^*\psi\,dx = c_1^*c_1\int|\phi_1|^2\,dx + c_1^*c_2e^{(i/\hbar)(E_1-E_2)t}\int\phi_1^*\phi_2\,dx \\ + c_1c_2^*e^{-(i/\hbar)(E_1-E_2)t}\int\phi_1\phi_2^*\,dx + c_2^*c_2\int\phi_2^*\phi_2\,dx\end{aligned} \tag{4-45}$$

Since this must give a constant result, the time-variable terms (i.e., terms including $e^{\pm(i/\hbar)(E_1-E_2)t}$) must vanish for all possible choices of c_1

and c_2. This means

$$\int_{-\infty}^{\infty} \phi_1^*\phi_2\,dx = \int_{-\infty}^{\infty} \phi_1\phi_2^*\,dx = 0 \tag{4-46}$$

Where two functions f, g satisfy $\int f^*g\,dx = 0$, we say they are orthogonal. Thus Eq. (4-46) says that two states of different energy are orthogonal.

Later we shall learn an interpretation for expressions such as $\int f^*g\,dx$, and we shall find that Eq. (4-46) records the fact that if a particle is known to have an energy E_1 (and hence a wave function $\psi_1 = e^{iE_1t/\hbar}\phi_1$), then the amplitude that it is found to have a different energy E_2 (i.e., wave function $e^{iE_2t/\hbar}\phi_2$) must be 0.

Problem 4-8 Show from the fact that H is hermitian that E is real. [Choose $f = g = \phi$ in Eq. (4-30).]

Problem 4-9 Show from the fact that H is hermitian that Eq. (4-46) holds. [Choose $f = \phi_2$, $g = \phi_1$ in Eq. (4-30).]

Linear Combinations of Steady-state Functions. Suppose that our functions corresponding to the set of energy levels E_n are not only orthogonal but also normalized, i.e., that the integral of the absolute square over all x is 1. Then we shall have

$$\int_{-\infty}^{\infty} \phi_n^*(x)\phi_m(x)\,dx = \delta_{nm} \tag{4-47}$$

where δ_{nm}, the Kronecker delta, is defined by $\delta_{nm} = 0$ if $n \neq m$ and $\delta_{nn} = 1$. Many functions can be expressed as a linear combination of such ϕ_n's. In particular, any function which is likely to arise as a wave function can be so expressed. That is,

$$f(x) = \sum_{n=1}^{\infty} a_n\phi_n(x) \tag{4-48}$$

The coefficients a_n are easily obtained; for by multiplication of Eq. (4-48) by the quantity $\phi_m^*(x)$ and integration over all x we obtain

$$\int_{-\infty}^{\infty} \phi_m^* f(x)\,dx = \sum_{n=1}^{\infty} a_n \int_{-\infty}^{\infty} \phi_m^*\phi_n\,dx = a_m \tag{4-49}$$

That is,

$$a_n = \int_{-\infty}^{\infty} \phi_n^*(x) f(x)\,dx \tag{4-50}$$

Thus we have the identity

$$f(x) = \sum_{n=1}^{\infty} \phi_n(x) \int_{-\infty}^{\infty} \phi_n^*(y) f(y)\, dy$$

$$= \int_{-\infty}^{\infty} \left[\sum_{n=1}^{\infty} \phi_n(x) \phi_n^*(y) \right] f(y)\, dy \tag{4-51}$$

An interesting alternative way of expressing this result makes use of the definition of the delta function. We write

$$\delta(x - y) = \sum_{n=1}^{\infty} \phi_n(x) \phi_n^*(y) \tag{4-52}$$

It is possible to express the kernel K in terms of these functions ϕ_n and the energy values E_n. We do so by the following consideration. Let us ask this: If $f(x)$ is the known wave function at the time t_1, what is the wave function at the time t_2? It can be written at any time t as

$$\psi(x,t) = \sum_{n=1}^{\infty} c_n e^{-(i/\hbar) E_n t} \phi_n(x) \tag{4-53}$$

for it is a solution of the Schrödinger equation, and any solution can be written in this form. But at the time t_1 we have

$$f(x) = \psi(x,t_1) = \sum_{n=1}^{\infty} c_n e^{-(i/\hbar) E_n t_1} \phi_n(x) = \sum_{n=1}^{\infty} a_n \phi_n(x) \tag{4-54}$$

since we can always express $f(x)$ in the form of Eq. (4-48). So we conclude

$$c_n = a_n e^{+(i/\hbar) E_n t_1} \tag{4-55}$$

Putting this into Eq. (4-53), we have

$$\psi(x,t_2) = \sum_{n=1}^{\infty} c_n e^{-(i/\hbar) E_n t_2} \phi_n(x) = \sum_{n-1}^{\infty} a_n e^{+(i/\hbar) E(t_1 - t_2)} \phi_n(x) \tag{4-56}$$

Now using Eq. (4-50) for the coefficient a_n, we obtain

$$\psi(x,t_2) = \sum_{n=1}^{\infty} \phi_n(x) e^{-(i/\hbar) E_n (t_2 - t_1)} \int_{-\infty}^{\infty} \phi_n^*(y) f(y)\, dy$$

$$= \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \phi_n(x) \phi_n^*(y) e^{-(i/\hbar) E_n (t_2 - t_1)} f(y)\, dy \tag{4-57}$$

This final expression determines the wave function at the time t_2 completely in terms of $f(x)$, the wave function at the time t_1. Previously we represented this relation by the equation

$$\psi(x,t_2) = \int_{-\infty}^{\infty} K(x,t_2;y,t_1)f(y)\,dy \tag{4-58}$$

Comparing this to Eq. (4-57), we finally obtain the desired expression for the kernel $K(2,1)$,

$$\begin{aligned} K(x_2,t_2;x_1,t_1) &= \sum_{n=1}^{\infty} \phi_n(x_2)\phi_n^*(x_1)e^{-(i/\hbar)E_n(t_2-t_1)} \qquad \text{for } t_2 > t_1 \\ &= 0 \qquad \text{for } t_2 < t_1 \end{aligned} \tag{4-59}$$

Problem 4-10 Verify that K as defined in Eq. (4-59) satisfies the Schrödinger equation (4-29).

This expression for K is very useful for translating expressions to more conventional representations. It expresses the kernel, which was originally a path integral, entirely in terms of solutions of the differential equation (4-42).

Problem 4-11 Show that for free particles in three dimensions the solutions

$$\phi_p = e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}} \tag{4-60}$$

go with the energy $E_p = p^2/2m$. Consider the vector $\mathbf{p}$ as an index n and note the orthogonality. That is, as long as $\mathbf{p} \neq \mathbf{p}'$,

$$\int^{\mathbf{r}} \phi_p^*\phi_{p'}\,d^3\mathbf{r} = 0 \qquad \text{even if } E_p = E_{p'} \tag{4-61}$$

Therefore the free-particle kernel must be

$$K_0(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1) = \sum_{\mathbf{p}} e^{-(i/\hbar)\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)}e^{-ip^2(t_2-t_1)/2\hbar m} \tag{4-62}$$

Since the $\mathbf{p}$'s are distributed over a continuum, the sum over the "indices" $\mathbf{p}$ is really equivalent to an integral over the values of $\mathbf{p}$, namely,

$$\sum_{\mathbf{p}}(\quad) = \int^{\mathbf{p}}(\quad)\frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \tag{4-63}$$

Therefore, we find that the free-particle kernel is given by

$$K_0(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1) = \int^{\mathbf{p}} e^{-(i/\hbar)\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)}e^{-ip^2(t_2-t_1)/2m\hbar}\frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \tag{4-64}$$

Problem 4-12 Carry out the integral in Eq. (4-64) by completing the square. Show that the form of the free-particle kernel results [i.e., the three-dimensional version of Eq. (3-3)].

4-3 NORMALIZING THE FREE - PARTICLE WAVE FUNCTIONS

The derivation of the kernel for a free particle, as given in Prob. 4-11, is unsatisfactory for two related reasons. First, the idea of a sum over distinct states n used in Eq. (4-62) is not satisfactory if the states lie in a continuum, as they do for a free particle where any p is allowed. Second, the plane-wave functions for free particles, although orthogonal, cannot be normalized, that is, $\int_{-\infty}^{\infty} \phi^*\phi\, dx = \int_{-\infty}^{\infty} 1\, dx = \infty$, so the condition of Eq. (4-47) used in deriving Eq. (4-62) is not satisfied. Both of these points can be remedied together in a perfectly straightforward mathematical way. Starting all the way back when we expressed an arbitrary function as a sum of eigenfunctions,

$$f(x) = \sum_n a_n\phi_n(x) \tag{4-65}$$

we allow part, or all, of the states to lie in a continuum, so that the sum on n must be replaced partly by an integral. With mathematical care one can find the correct expression for K analogous to Eq. (4-62) but applying also when the states are in a continuum.

Normalizing in a Box. Many physicists prefer another, less rigorous approach. What they do is modify the original problem in a way that (from physical reasoning) will not essentially modify the result yet will leave all the states separate in energy and all the simple sums as simple sums. In our example this may be accomplished as follows. We are studying the amplitude that in a finite time a particle goes from x_1 at time t_1 to x_2 at time t_2. Now if these two points are some finite distance apart and the time is not extremely long, surely it can make no appreciable difference to the amplitude whether the electron is really free or is instead confined to some enormous box of volume V with walls very, very far from x_1 and x_2. The amplitude could be affected only if the particle could run out to the walls and back in the time $t_2 - t_1$; but if the walls are far enough away, there is no appreciable amplitude for this.

It is always possible that this assumption fails for some special-shaped walls such that, for example, x_2 is at a focus of waves from x_1

reflected at the walls. From time to time someone lets an error creep in by replacing a system in empty space with one at the center of a large spherical box. The fact that the system remains at the exact center of a perfect sphere may have an effect (like the spot of light at the center of the shadow of a perfectly circular object) which does not vanish as the sphere radius goes to infinity. For another shape, or a system off-center to the sphere, the surface effect would vanish.

Take first the case of one dimension. In empty space the space-dependent wave functions are $e^{ipx/\hbar}$ (any p, positive or negative). If, instead, the range of x is limited to $-L/2$ to $+L/2$, say, what are the functions ϕ? The answer depends on the boundary conditions defining ϕ at $x = -L/2$ and $x = +L/2$. The easiest conditions to understand physically are those for walls which offer very high repulsive potentials to the particle, thus confining it (i.e., perfect reflectors). They correspond to $\phi(x) = 0$ at $x = -L/2$ and $x = +L/2$. The solutions of the wave equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2} = E\phi \tag{4-66}$$

in the range $|x| < L/2$ are, for $E = p^2/2m = \hbar^2k^2/2m$,

$$e^{ikx} \qquad \text{and} \qquad e^{-ikx}$$

or any linear combination. Neither e^{ikx} nor e^{-ikx} can satisfy the boundary conditions, but with $k = n\pi/L$ (n an integer) satisfactory solutions are given by half the sum (which is $\cos kx$) for n odd and $\frac{1}{2}i$ times the difference (which is $\sin kx$) for n even, as diagramed in Fig. 4-1. Thus the states are sines and cosines and the energy levels are separated (i.e., not in a continuum).

If the solutions are written as

$$\sqrt{\frac{2}{L}}\cos kx \qquad \text{and} \qquad \sqrt{\frac{2}{L}}\sin kx$$

they are normalized, since

$$\int_{-L/2}^{L/2}\left(\sqrt{\frac{2}{L}}\cos kx\right)^2 dx = 1 \tag{4-67}$$

A sum over states is a sum over n. If we consider, say, the sine wave functions (thus, even values of n) for very large L but not large x (walls far from the point of interest), the successive functions differ

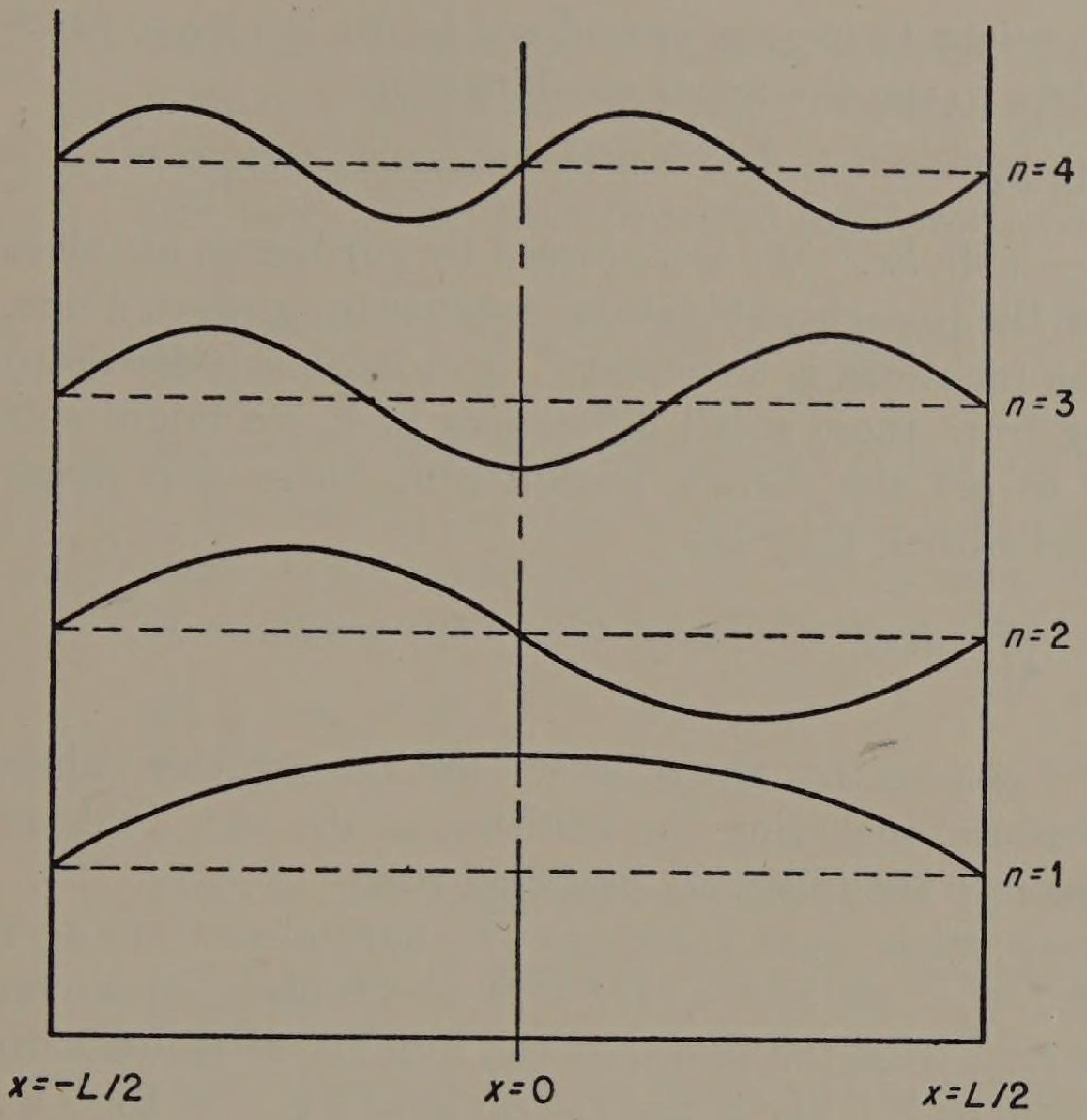


Fig. 4-1 The form of the one-dimensional wave functions which have been normalized in a box. The first four are shown. The corresponding energy levels are $E_1 = \hbar^2\pi^2/2mL^2$, $E_2 = 4E_1$, $E_3 = 9E_1$, and $E_4 = 16E_1$. The absolute magnitude of the energy, which depends on the size of our fictitious box, is not important for more realistic problems. Rather, it is the relation between the energy levels of the various states which has significance.

by only a small amount. This difference

$$\sqrt{\frac{2}{L}}\left[\sin 2\pi(n+1)\frac{x}{L} - \sin 2\pi n\frac{x}{L}\right] = 2\sqrt{\frac{2}{L}}\cos 2\pi\frac{2n+1}{2}\frac{x}{L}\sin 2\pi\frac{x}{2L}$$

$$\simeq \sqrt{\frac{2}{L}}\frac{2\pi x}{L}\cos 2\pi\left(n+\frac{1}{2}\right)\frac{x}{L} \qquad (4\text{-}68)$$

is approximately proportional to the small quantity x/L. So a sum on n can be replaced by an integral over $k = 2\pi n/L$. Since the successive allowed values of k (for sine functions) are spaced by $2\pi/L$, there are $\frac{L}{2\pi}\Delta k$ states in range Δk. All of this applies also to states with the cosine wave function, so that we may replace sums by integrals in our formulas with

$$\sum_{n=0}^{\infty}(\quad) \to \int_{\infty}^{0}(\quad)\frac{dk}{2\pi}L \qquad (4\text{-}69)$$

and remember to add the result for the two kinds of wave functions, namely, $\sqrt{2/L}\cos kx$ and $\sqrt{2L}\sin kx$.

It is often inconvenient to use sin kx and cos kx for the wave functions as we would like to use the linear combinations

$$e^{ikx} = \cos kx + i \sin kx$$

and $e^{-ikx} = \cos kx - i \sin kx$. We were forced by our box to use sines and cosines and not the linear combination, because for a given k one, but not both, of the functions is a solution. But if we can disregard small errors arising from these small differences in k, we might still expect to be able to get the correct results from these new linear combinations. Normalized, they are

$$\sqrt{\frac{1}{L}}\, e^{ikx} \qquad \text{and} \qquad \sqrt{\frac{1}{L}}\, e^{-ikx}$$

Since the wave e^{-ikx} can be thought of as e^{ikx} but for negative values of k, our new procedure, including the addition of the two kinds of wave functions, becomes the following practical rule:

To deal with free-particle wave functions e^{ikx}, normalize them to a range of x of length L (i.e., use $\phi = \sqrt{1/L}\, e^{ikx}$), and replace sums over states by integrals on k with the rule that the number of states with k in the range k to $k + dk$ is $L\, dk/2\pi$ and the range of k is $-\infty$ to $+\infty$.

Periodic Boundary Conditions. Sometimes this excursion into cosines and sines and back to exponentials is avoided by the following argument. The wall is artificial anyway, so its particular position and the particular boundary condition should not make any physical difference as long as it is far away. So instead of the physically simple conditions $\phi = 0$ at $x = L/2$ and at $x = -L/2$, let us use two others for which the solutions are indeed e^{ikx} directly. These are

$$\phi(x) \quad \left(\text{at } x = \frac{L}{2}\right) = \phi(x) \quad \left(\text{at } x = \frac{-L}{2}\right) \tag{4-70}$$

and

$$\phi'(x) \quad \left(\text{at } x = \frac{L}{2}\right) = \phi'(x) \quad \left(\text{at } x = \frac{-L}{2}\right) \tag{4-71}$$

These are called *periodic boundary conditions*, because the same ones would result by the requirement that ϕ is periodic in x in all space with period $x = L$. It is readily verified that the functions $\sqrt{1/L}\, e^{ikx}$ are solutions, normalized to range L, provided $k = 2\pi n/L$ with n an integer, positive, negative, or zero. From this our rule follows directly.

In three dimensions we can see what happens by using a rectangular box of sides L_x, L_y, L_z in the three directions. Let us use periodic boundary conditions. That is, the magnitude and first derivative of a wave function at a point on one face are respectively equal to the magnitude and first derivative at the corresponding point on the opposite face. The wave function for a free particle normalized is

$$\sqrt{\frac{1}{L_x}}\,e^{ik_xx}\sqrt{\frac{1}{L_y}}\,e^{ik_yy}\sqrt{\frac{1}{L_z}}\,e^{ik_zz} = \frac{1}{\sqrt{V}}\,e^{i\mathbf{k}\cdot\mathbf{r}} \tag{4-72}$$

where $V = L_xL_yL_z$ is the volume of the box and the allowed values of k_x are $2\pi n_x/L_x$ for n_x an integer, those of k_y are $2\pi n_y/L_y$ for n_y an integer, and those of k_z are $2\pi n_z/L_z$ for n_z an integer. Furthermore, the number of solutions with k_x in range dk_x, k_y in dk_y, and k_z in dk_z is

$$\frac{dk_x}{2\pi}L_x\frac{dk_y}{2\pi}L_y\frac{dk_z}{2\pi}L_z = \frac{d^3\mathbf{k}}{(2\pi)^3}V \tag{4-73}$$

That is, use plane waves normalized to volume V: $\sqrt{\frac{1}{V}}\exp(i\mathbf{k}\cdot\mathbf{r})$. The number of states in range $d^3\mathbf{k}$ (differential volume of $\mathbf{k}$ space) is $V\,d^3\mathbf{k}/(2\pi)^3$.

Let us apply this to Prob. 4-11 and recall the connection between momentum and wave number $p = \hbar k$ brought out in Sec. 3-1. In Eq. (4-64) we must make two changes. First, since the wave functions used were $\exp(i\mathbf{p}\cdot\mathbf{r}/\hbar)$, whereas we should have used

$$\sqrt{\frac{1}{V}}\exp\left(\frac{i\mathbf{p}\cdot\mathbf{r}}{\hbar}\right)$$

there should be an additional factor $1/V$; for the product of two wave functions was involved. Second, the symbol $\sum_{\mathbf{p}}(\quad)$ must be replaced by $V\int(\quad)\,d^3\mathbf{p}/(2\pi\hbar)^3$. This justifies what was done in Sec. 4-2 and the results given in Prob. 4-11.

It is noted that the V factors cancel out, as indeed they must; for as $V \to \infty$ the kernel K must be independent of the size of the box.

Some Remarks on Mathematical Rigor. The reader may have one or two reactions on seeing how the volumes V cancel at the end of this calculation. One might be: How nicely it cancels out as it should, for the walls have no effect. The other might be: Why do it in this complicated and "dirty" nonrigorous manner, putting in walls which make no difference, etc., when all this can be done much more elegantly and rigorously mathematically without the need of walls, etc? It depends whether you are physically minded or mathematically minded. There are many misunderstandings between mathematicians and physicists on the place of mathematical rigor in physics, so perhaps a word as to the value of each method (the box or mathematical rigor) may be in order.

There is, of course, the more trivial point: Which is most familiar—which takes the least new knowledge? Most physicists have seen this argument about how to count the states in a box before.

Another point is that the mathematically rigorous solution may not be physically rigorous. That is, the box may in fact exist. It may not be a rectangular box, but it is not often that experiments are done under the stars. Rather they are done in a room. Although it is physically reasonable that the walls have no effect, it is true that the original problem is set up as an idealization. It is no more satisfactory idealization to move the walls to infinity than to replace them by perfect mirrors far away. The mathematical rigor is wasted in the first idealization, since the walls are not at infinity.

The wall approach is just as rigorous, or rather just as nonrigorous. It has several advantages. For example, in finding that the volume cancels out we do learn that at least one aspect of the idealized walls, namely how far away they are, is unimportant. This discovery makes us more intuitively convinced that the actual disposition of the real environment may be unimportant. Finally, the formula derived is very useful when in fact we do have a finite sample. For example, in Chap. 8 we shall use it to count sound-wave modes in a large, rectangular block of material.

On the other hand, the advantage of the mathematically clean argument is the avoidance of much unnecessary detail that cancels out. Although, using the wall approach, one may learn something about how the walls have no effect, one may be firmly convinced that this is true anyway and not wish to descend into details to see it again.

The normalization problem is a special example, but it illustrates the point. The physicist cannot understand the mathematician's care in solving an idealized physical problem. The physicist knows the real problem is much more complicated. It has already been simplified by intuition, which discards the unimportant and often approximates the remainder.

5

Measurements and Operators

So far we have described quantum-mechanical systems as if we intended to measure only the coordinates of position and time. Indeed, all measurements of quantum-mechanical systems could be made to reduce eventually to position and time measurements (e.g., the position of a needle on a meter or the time of flight of a particle). Because of this possibility a theory formulated in terms of position measurements is complete enough in principle to describe all phenomena. Nevertheless, it is convenient to try to answer directly a question involving, say, a measurement of momentum without insisting that the ultimate recording of the equipment must be a position measurement and without having to analyze in detail that part of the apparatus which converts momentum to a recorded position. Thus, in this chapter, instead of concentrating on the amplitude that a particle has a definite position, we shall develop the idea of an amplitude to find a definite momentum, energy, or other physical quantity.

In the first section of the chapter we shall show how a system may be described in terms of momentum and energy. The concepts learned here will be extended in the second section to describe in general various ways of representing the quantum-mechanical system. The transformation functions which enable us to go from one method of representation to another have many interesting properties. Among them is the concept of an operator, which was introduced in the preceding chapter and will be discussed further in the third section of this chapter.

5-1 THE MOMENTUM REPRESENTATION

The Momentum Amplitude. So far we have used the concept of probability in terms of the position of a particle, but suppose we wish to measure the momentum. Is there an amplitude $\phi(p)$ whose absolute square will give us the probability $P(p)$ that a measurement of momentum will show that the particle has the momentum p? There is in fact such an amplitude, and we can easily find it.

Some ways of measuring momentum (or other physical quantities) correspond to measurements of position, and thus they can be analyzed if we know how to analyze coordinate measurements. For example, working in one dimension, suppose we have a particle whose position at $t = 0$ is localized within $\pm b$ of the origin of the x axis. The uncertainty b can be as large as desired so long as it is finite. We can measure the momentum of such a particle by a time-of-flight technique. That is, we can observe how far the particle has traveled (assuming no forces) by the time $t = T$. If the position is x, then the velocity is x/T and the momentum is $p = mx/T$. The error in such a momentum measurement, $\pm mb/T$, can be made as small as desired by making T sufficiently large.

Suppose we analyze the momentum probability $P(p)$ as defined by such an experiment. The probability $P(p)\,dp$ that the momentum lies between p and $p + dp$ is the probability $P(x)\,dx$ that, if all the potentials affecting the particle are suddenly turned off, then after the time T the particle will be found between the points x and $x + dx$. Of course, this requires that we connect p with x by $p = mx/T$. Assume the wave function of the particle as given by $f(y)$ at $t = 0$, and our problem is to find $P(p)$ directly in terms of $f(y)$.

The amplitude for the particle to arrive at x at the time $t = T$ is

$$\psi(x,T) = \int_{-\infty}^{\infty} K_0(x,T;y,0)f(y)\,dy \tag{5-1}$$

Upon substitution for the free-particle kernel K_0, this expression becomes

$$\psi(x,T) = \sqrt{\frac{m}{2\pi i\hbar T}}\exp\frac{imx^2}{2\hbar T}\int_{-\infty}^{\infty}\left(\exp\frac{-imxy}{\hbar T}\right)\left(\exp\frac{imy^2}{2\hbar T}\right)f(y)\,dy \tag{5-2}$$

The absolute square of this amplitude gives the probability that the particle lies between x and $x + dx$. According to our definition, this is identical (in the limit as $T \to \infty$) with the probability that the momentum of the particle lies between p and $p + dp$.

$$\begin{aligned} P(x)\,dx &= \frac{m\,dx}{2\pi\hbar T}\left|\int_{-\infty}^{\infty}\left\{\exp\left[\frac{im}{2\hbar T}(y^2 - 2xy)\right]\right\}f(y)\,dy\right|^2 \\ &= P(p)\,dp \qquad \text{as } T \to \infty \end{aligned} \tag{5-3}$$

Then substituting $p = mx/T$, and supposing that we pass to the limit of large T, there results

$$P(p)\,dp = \frac{dp}{2\pi\hbar}\left|\int_{-\infty}^{\infty}\left[\exp\left(\frac{imy^2}{2\hbar T} - \frac{ipy}{\hbar}\right)\right]f(y)\,dy\right|^2 \tag{5-4}$$

We assumed earlier that, initially, the particle would be restricted to a region within $\pm b$ of the origin. This means that the initial wave function $f(y)$ drops to 0 for values of y larger in absolute magnitude than b. Now as T becomes large the quantity $imb^2/2\hbar T$ becomes negligibly small. Since there is no contribution to the integral of Eq.

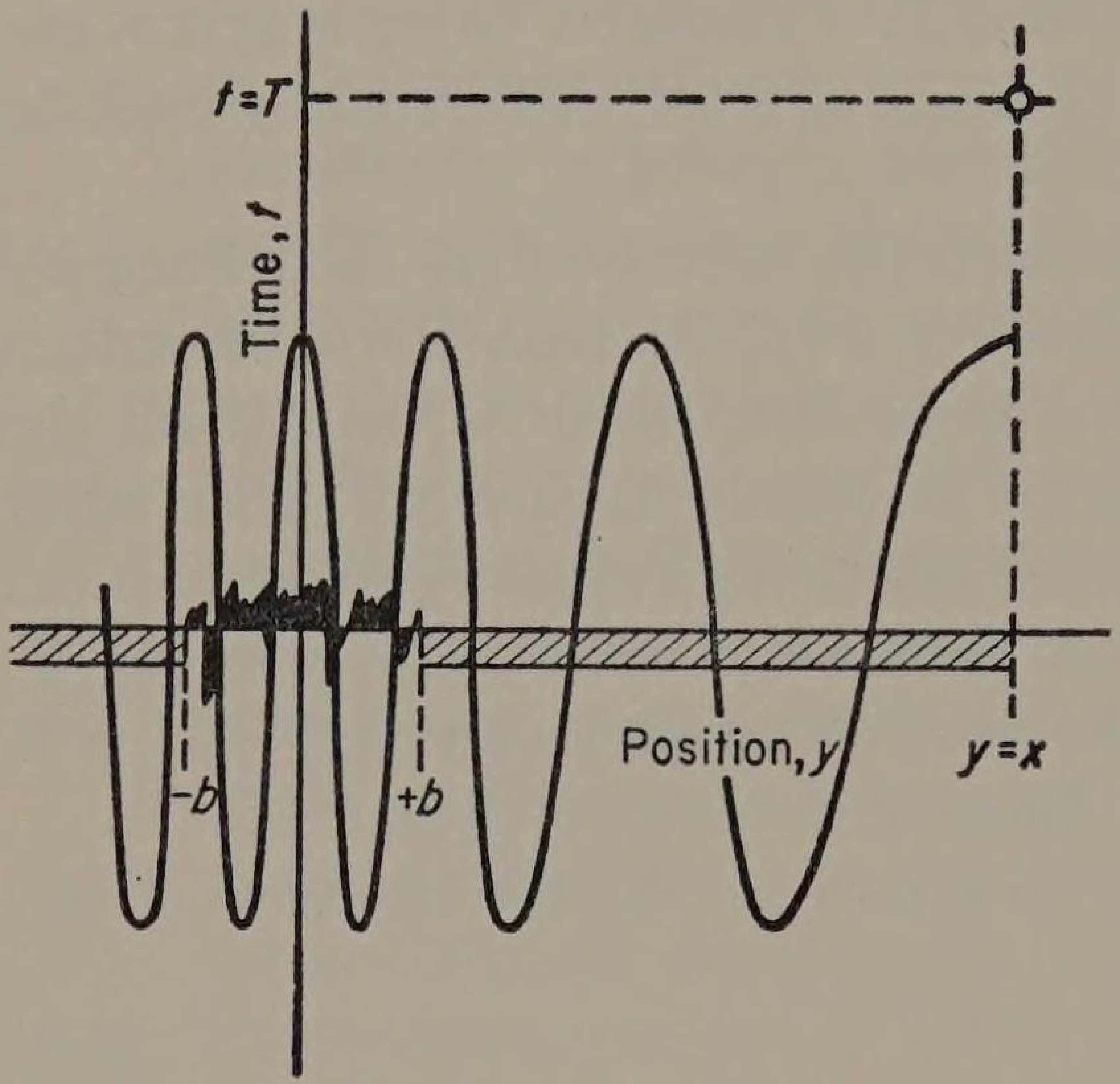


Fig. 5-1 The amplitude for particles traveling freely to arrive at the position x in the time interval T is determined by the product of two functions. The first is the amplitude $f(y)$ for the particle to start at some position y, as shown by the shaded curve in the figure. The second, the amplitude to get from y to x, is the free-particle kernel $K(x,T;y,0)$, as shown by the sine wave of slowly changing wavelength in the figure. (Here we treat the final position x as the origin of this function, since y is our variable.) If the point x is far from the origin, compared to the distance $-b$ to $+b$ over which $f(y)$ is nonzero, the wave has an approximately constant wavelength. Its approximate form is $\exp[(-i/\hbar)(mx/T)y]$. The two functions are multiplied together and then integrated over y to get the final amplitude for arrival. Since all particles have traveled approximately the same distance x (again assuming $x \gg b$) in the same time T, this final amplitude is equivalent to the amplitude that the particles have the momentum $p = mx/T$.

(5-4) for values of y greater in absolute magnitude than b, the probability $P(p)\,dp$ approaches $dp/2\pi\hbar$ times the absolute square of the amplitude[1]

$$\phi(p) = \int_{-\infty}^{\infty} \left(\exp\frac{-ipy}{\hbar}\right) f(y)\,dy \tag{5-5}$$

An alternative explanation of this result is given in Fig. 5-1 and extended in Fig. 5-2.

The expression for the momentum amplitude given by Eq. (5-5) applies to a one-dimensional situation. It is easy to extend the defi-

---

[1] Many writers prefer to account for the factor $1/2\pi\hbar$ in the definition of $\phi(p)$, where it appears as $1/\sqrt{2\pi\hbar}$. However, following the development of Sec. 4-3, we prefer to write it in the form we have used and remember that the differential element of momentum always includes the factor $1/2\pi\hbar$ in each dimension. For example, the differential element of momentum in three-dimensional momentum space is $d^3\mathbf{p}/(2\pi\hbar)^3$.

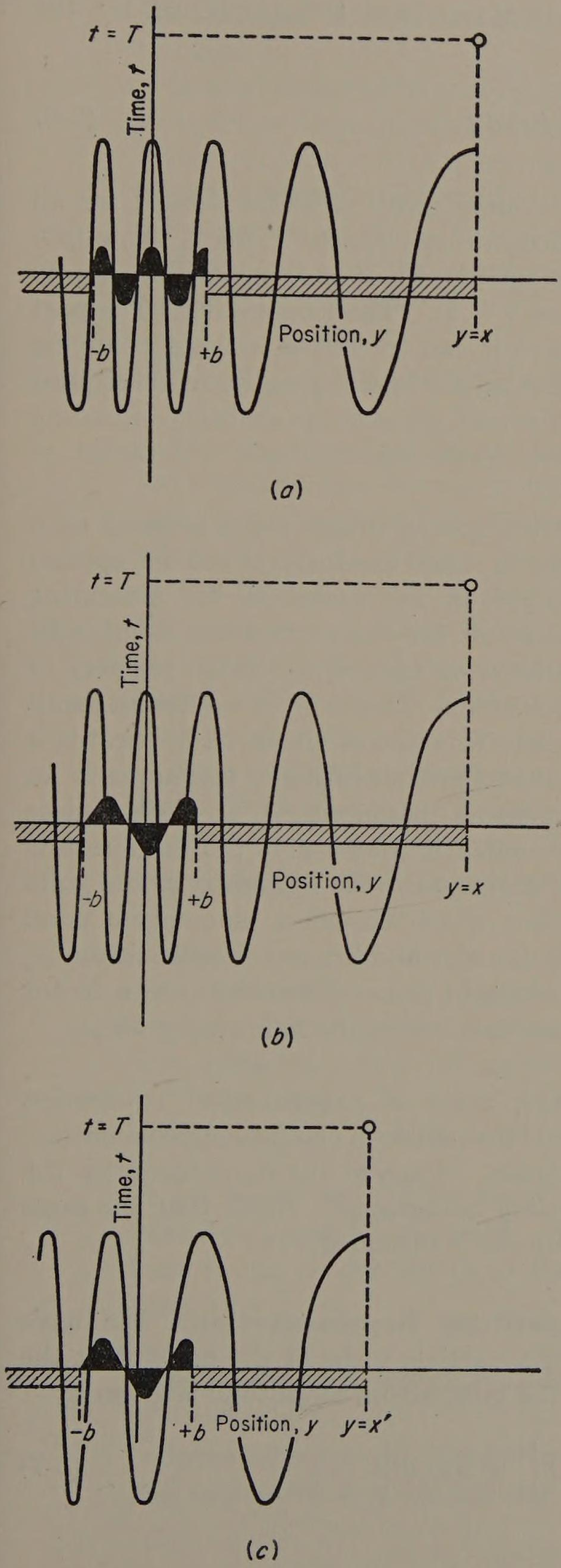


Fig. 5-2 If the amplitude $f(y)$ is roughly periodic with the same wavelength as the overlying kernel, as shown in (*a*), then the integral of the product of the two functions is very large. That is, the probability that the momentum is $p = mx/T$ is large.

On the other hand, suppose the wavelengths differ for some new function $f'(y)$ as shown in (*b*). Then, when the product is taken, the contributions to the integral from different values of y will cancel each other. Now the probability that the momentum is mx/T is small.

If a new position x' is chosen as a final point, as shown in (*c*), then a new region of the kernel curve overlies the space $-b$ to $+b$. For a correct choice of x', the wavelength of the kernel in this new region is the same as the wavelength of $f'(y)$ and a large probability results. That is, there is a large probability that such particles have the new momentum value $p = mx'/T$.

nition to the three-dimensional case where the amplitude for the momentum is

$$\phi(\mathbf{p}) = \int^{\mathbf{r}} \left\{\exp\left[-\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r})\right]\right\} f(\mathbf{r})\, d^3\mathbf{r} \tag{5-6}$$

where the wave function $f(\mathbf{r})$ is now assumed to be defined for all points in the three-dimensional coordinate space. This is the amplitude that the particle has the momentum $\mathbf{p}$ at the time $t = 0$. (Note that it is *not* defined for the time $t = T$. The time interval T is part of the measuring equipment, and it can be varied without changing the momentum amplitude.) The square of this amplitude, multiplied by the differential momentum element, gives the probability of finding the momentum in the interval (three-dimensional) $d^3\mathbf{p}/(2\pi\hbar)^3$ of momentum space.

We have analyzed a momentum measurement which is based on a time-of-flight technique. However, such an analysis can be applied to other techniques. The analysis of any technique for measuring momentum will give the same result for the momentum amplitude. For suppose we have two methods or techniques which purport to measure the same quantity, momentum. If one gives a different result than the other, we have to explain why one or the other apparatus is faulty. So if you will grant that the time-of-flight technique is an adequate way to define a momentum measurement, any other piece of equipment which measures momentum must give the same results $P(p)\,dp$ for the distribution of momenta if the system is in the state $f(y)$. Analysis of any equipment which measures momentum must give the same expression $\phi(p)$ for the amplitude for momentum p, within possibly an irrelevant constant phase difference (i.e., a factor $e^{i\delta}$ with δ constant). For example, consider the following problem.

Problem 5-1 Consider any piece of experimental equipment designed to measure momentum by means of a classical approximation, such as a magnetic field analyzer. Analyze the equipment by the methods outlined in the preceding paragraphs. Show that the same result for the momentum amplitude is obtained.

Transformation to Momentum Representation. We have called $\psi(\mathbf{R},t)$ the amplitude for a particle to be at the point $\mathbf{R}$ at the time t. We have found that the momentum amplitude is given by

$$\phi(\mathbf{p},t) = \int^{\mathbf{R}} \left\{\exp\left[-\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{R})\right]\right\} \psi(\mathbf{R},t)\, d^3\mathbf{R} \tag{5-7}$$

We shall call this the amplitude that the particle has momentum $\mathbf{p}$ at the time t.

It is often useful to analyze problems in this momentum representation rather than in the coordinate representation, or, as it is often stated, in momentum space rather than in coordinate space. Actually, the transformation from one representation to the other is just a Fourier transform. Thus if we have the momentum representation and wish to find the coordinate representation, we use the inverse transform given by

$$\psi(\mathbf{R},t) = \int^{\mathbf{p}} \left[\exp \frac{i}{\hbar} (\mathbf{p} \cdot \mathbf{R}) \right] \phi(\mathbf{p},t) \frac{d^3\mathbf{p}}{(2\pi\hbar)^3} \tag{5-8}$$

We can describe this last formula in the same physical terms we have used to describe the structure of other amplitudes. The amplitude that the particle is at the position $\mathbf{R}$ is given by the sum over alternatives. In this case the alternatives correspond to the products of two terms. One of these is the amplitude that the momentum of the particle is $\mathbf{p}$, given by $\phi(\mathbf{p})$. The other term, $\exp(i\mathbf{p} \cdot \mathbf{R}/\hbar)$, *is the amplitude that if the momentum is* $\mathbf{p}$, *then the particle is at the position* $\mathbf{R}$. This second factor is not new to us, for we have discussed such an expression in Prob. 3-4.

Note that in the transform of Eq. (5-7) the exponent has a minus sign. Such a term can be described in a manner parallel to that used in the preceding paragraph. Thus we can say that $\exp(-i\mathbf{p} \cdot \mathbf{R}/\hbar)$ *is the amplitude that if a particle is at position* $\mathbf{R}$, *it has the momentum* $\mathbf{p}$.

The Kernel in Momentum Representation. We have shown (Sec. 3-4) how a wave function at a particular time t_2 can be derived from the wave function at an earlier time t_1 with the help of the kernel describing the motion of the particle in the intervening time. Thus

$$\psi(\mathbf{R}_2,t_2) = \int_{t_1}^{t_2} \int^{\mathbf{R}_1} K(\mathbf{R}_2,t_2;\mathbf{R}_1,t_1)\psi(\mathbf{R}_1,t_1)\, d^3\mathbf{R}_1\, dt \tag{5-9}$$

It is possible to define a kernel in momentum space which would be used in a parallel expression. Thus the momentum amplitude at the time t_2 can be derived from the momentum amplitude at an earlier time t_1 by

$$\phi(\mathbf{p}_2,t_2) = \int_{t_1}^{t_2} \int^{\mathbf{p}_1} \mathcal{K}(\mathbf{p}_2,t_2;\mathbf{p}_1,t_1)\phi(\mathbf{p}_1,t_1) \frac{d^3\mathbf{p}_1}{(2\pi\hbar)^3}\, dt_1 \tag{5-10}$$

Substituting in Eq. (5-9) for $\psi(\mathbf{R}_1,t_1)$ the expression of Eq. (5-8) and taking the Fourier transform of $\psi(\mathbf{R}_2,t_2)$ to get $\phi(\mathbf{p}_2,t_2)$, as in Eq.

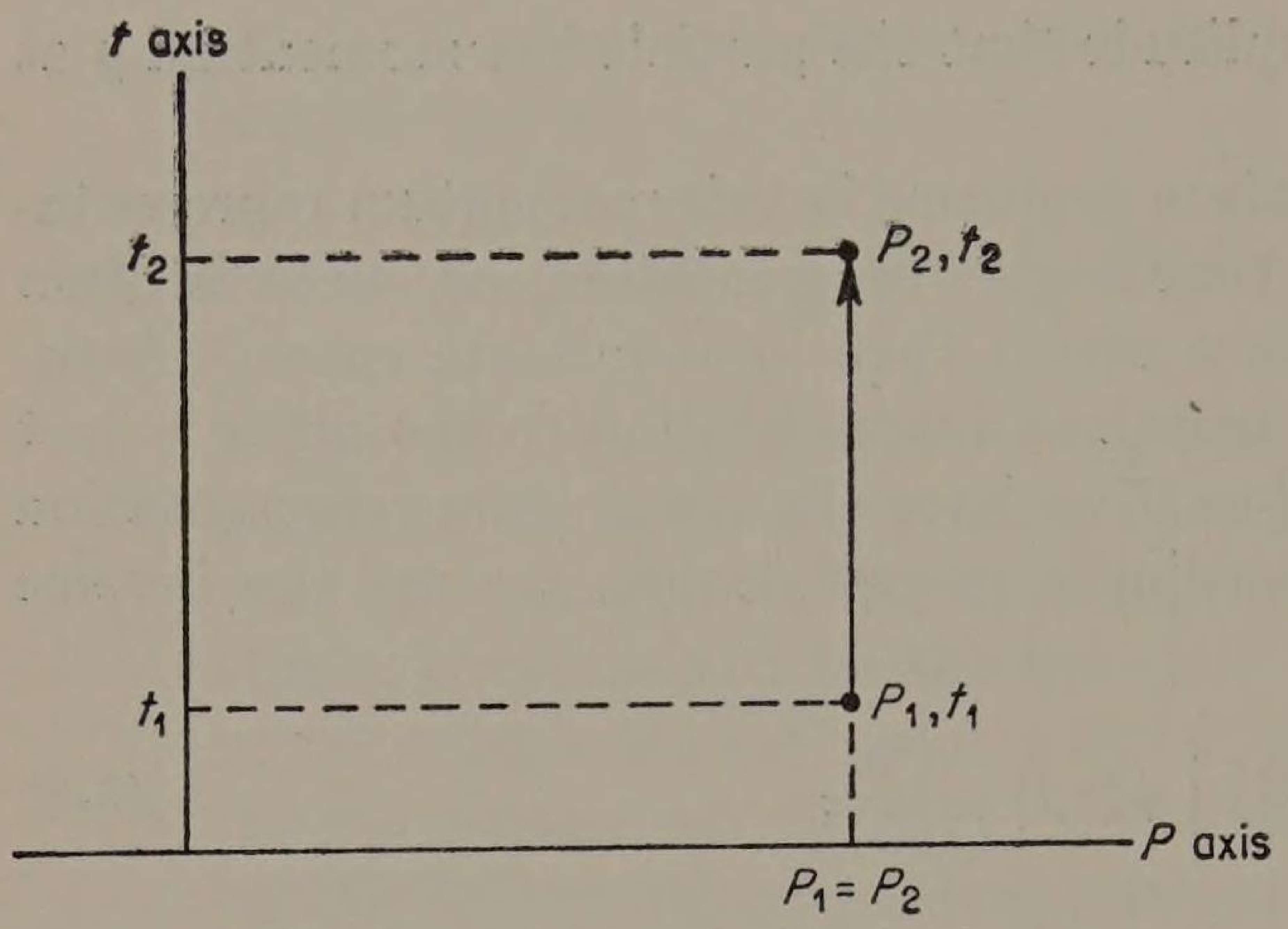


Fig. 5-3 The kernel for a free particle in momentum space is unlike the kernel in coordinate space. In momentum space, there is only one path which can carry the particle to the momentum value p_2 at the time t_2. That single path must start at the momentum $p_1 = p_2$. No other paths contribute to the kernel.

(5-7), we see that the kernel in momentum representation is given in terms of the kernel in coordinate representation by the expression

$$\mathcal{K}(\mathbf{p}_2,t_2;\mathbf{p}_1,t_1) = \int^{\mathbf{R}_1}\int^{\mathbf{R}_2} e^{-(i/\hbar)\mathbf{p}_2\cdot\mathbf{R}_2} K(\mathbf{R}_2,t_2;\mathbf{R}_1,t_1) e^{+(i/\hbar)\mathbf{p}_1\cdot\mathbf{R}_1} \, d^3\mathbf{R}_1 \, d^3\mathbf{R}_2 \tag{5-11}$$

For example, the kernel describing the motion of a free particle in momentum space is

$$\begin{aligned}\mathcal{K}_0(\mathbf{p}_2,t_2;\mathbf{p}_1,t_1) &= \int^{\mathbf{R}_1}\int^{\mathbf{R}_2} e^{-(i/\hbar)\mathbf{p}_2\cdot\mathbf{R}_2} K_0(\mathbf{R}_2,t_2;\mathbf{R}_1,t_1) e^{(i\hbar)\mathbf{p}_1\cdot\mathbf{R}_1} \, d^3\mathbf{R}_1 \, d^3\mathbf{R}_2 \\ &= (2\pi\hbar)^3\delta^3(\mathbf{p}_1 - \mathbf{p}_2) \exp\left[-\frac{i|\mathbf{p}_1|^2}{2\hbar m}(t_2 - t_1)\right] \qquad \text{for } t_2 > t_1 \\ &= 0 \qquad \text{for } t_2 < t_1\end{aligned} \tag{5-12}$$

[The last line follows the convention of Eq. (4-28).] The occurrence of the delta function in this expression shows that the momentum of a free particle does not change, as diagramed in Fig. 5-3. However, the phase of the momentum wave function changes continuously in accordance with the factor $e^{-iEt/\hbar}$, where $E = p^2/2m$. This result given by Eq. (5-12) can also be seen directly from Eq. (4-64).

This latter kernel offers a much simpler representation of the free particle than does the kernel in coordinate space. Generally, when the particle is not free, but rather moves under the influence of a potential, the kernel in momentum representation loses its simplicity. But if the effect of the potential can be represented in a perturbation expansion, this simplicity is regained (Chap. 6).

The Energy-Time Transformation. For many applications, particularly in relativistic quantum mechanics, it is best to treat the variables of space and time in a symmetric manner. Then in transforming from coordinate representation to momentum representation we include a transformation from time to energy. Thus the complete transformation for a kernel is

$$k(\mathbf{p}_2,E_2;\mathbf{p}_1,E_1) = \int^{\mathbf{R}_1}\int^{\mathbf{R}_2}\int_{-\infty}^{\infty}\int_{t_1}^{\infty} e^{-(i/\hbar)\mathbf{p}_2\cdot\mathbf{R}_2}e^{(i/\hbar)E_2t_2}K(\mathbf{R}_2,t_2;\mathbf{R}_1,t_1)$$

$$e^{(i/\hbar)\mathbf{p}_1\cdot\mathbf{R}_1}e^{-(i/\hbar)E_1t_1}\,d^3\mathbf{R}_1\,d^3\mathbf{R}_2\,dt_1\,dt_2 \qquad (5\text{-}13)$$

The energy E is not equal to $p^2/2m$, but is instead an extra independent variable (the coefficient of time) needed to define the kernel. Only if the system exists in the same energy state for an infinite time can an exact measurement of E be made to establish the relation between energy and momentum.

As an example, we shall work out the kernel for a free particle. For this case the integrals over $\mathbf{R}_1$ and $\mathbf{R}_2$ have already been worked out, with the results given in Eq. (5-12). Thus we are left with the integrals over t_1 and t_2. Make the substitution $t_2 = t_1 + \tau$. Then the double integral can be written as

$$\int_{-\infty}^{\infty} e^{(i/\hbar)(E_2-E_1)t_1}\,dt_1 \int_0^{\infty} e^{(i/\hbar)(E_2-p^2/2m)\tau}\,d\tau \qquad (5\text{-}14)$$

The first of these two integrals is a representation of the Dirac δ function. In particular it is $2\pi\hbar\delta(E_2 - E_1)$. The second integral is of the form

$$\int_0^{\infty} e^{i\omega\tau}\,d\tau \qquad (5\text{-}15)$$

This latter integral arises often in quantum-mechanical problems. If ω is a real number, the integral does not converge. In order to carry out the present calculation, we shall replace ω with a complex number $\omega + i\epsilon$. When both ω and ϵ are real numbers, the integral has the value $i/(\omega + i\epsilon)$.

Now it would be possible to take the limit of this fraction as ϵ approaches 0 and interpret the result simply as i/ω. However, such an interpretation would lead to incorrect (or rather, incomplete) results in further work. The function we are evaluating is a kernel, and in future work it will often be integrated (multiplied by some other function) over values of ω or its equivalent. If ϵ were dropped from the expression, then such integrals would have a pole at $\omega = 0$, and we would be at a loss what to do.

It would not be correct to take just the principal part of the integral at such a pole. This would give the wrong result. In particular, such a result would imply that the inverse transform of the kernel would not give back the original coordinate representation kernel with which we started. Such a transform would differ from the correct kernel in that it would not be zero for values of time less than zero. One way to obtain the correct result from such integrals is to place the pole an infinitesimal distance above the real axis. This is accomplished by leaving ϵ in the expression.

If we rationalize the expression as

$$\frac{i}{\omega + i\epsilon} = \frac{i(\omega - i\epsilon)}{\omega^2 + \epsilon^2} = \frac{i\omega}{\omega^2 + \epsilon^2} + \frac{\epsilon}{\omega^2 + \epsilon^2} \tag{5-16}$$

we can interpret the first term on the right-hand side as i/ω and in further integrations use the principal part of an integral involving this term. The second term becomes $\pi\delta(\omega)$ as ϵ approaches 0, and it is to be interpreted as such in further integrations. That is, if a more precise mathematical definition is wanted, $i/(\omega + i\epsilon)$ should be replaced by $\text{P.P.}(i/\omega) + \pi\delta(\omega)$. This means that

$$\begin{aligned}\int_0^\infty e^{i\omega\tau}\,d\tau &= \lim_{\epsilon\to 0} \frac{i}{\omega + i\epsilon}\\ &= \text{P.P.}\left(\frac{i}{\omega}\right) + \pi\delta(\omega)\end{aligned} \tag{5-17}$$

In all expressions containing ϵ, a limit at $\epsilon \to 0$ is implied.

Returning to the evaluation of the kernel, we replace ω with $E_2 - p^2/2m$ to find

$$k_0(\mathbf{p}_2,E_2;\mathbf{p}_1,E_1) = (2\pi\hbar)^4\delta^3(\mathbf{p}_2 - \mathbf{p}_1)\delta(E_2 - E_1)i\left(E_1 - \frac{p_1{}^2}{2m} + i\epsilon\right)^{-1} \tag{5-18}$$

The existence of the delta functions in this expression means that neither the energy E nor the momentum $\mathbf{p}$ changes during the motion of a free particle. These two quantities affect the motion of the particle as shown by the last term of this equation. That is, the amplitude for the motion from one point to another of a free particle with energy E and momentum $\mathbf{p}$ is proportional to $i(E - p^2/2m + i\epsilon)^{-1}$.

Earlier in this section it was mentioned that the energy E is not in general identical to $p^2/2m$, but is instead a separate variable. To understand the distinction, let us look at the kernel for a free particle, which is a wave-like function in time and space and wherein E is the coefficient of time and thus has the properties of a frequency. This kernel, given in Eq. (5-12), has the form shown in Fig. 5-4 when plotted against the time difference $T = t_2 -- t_1$.

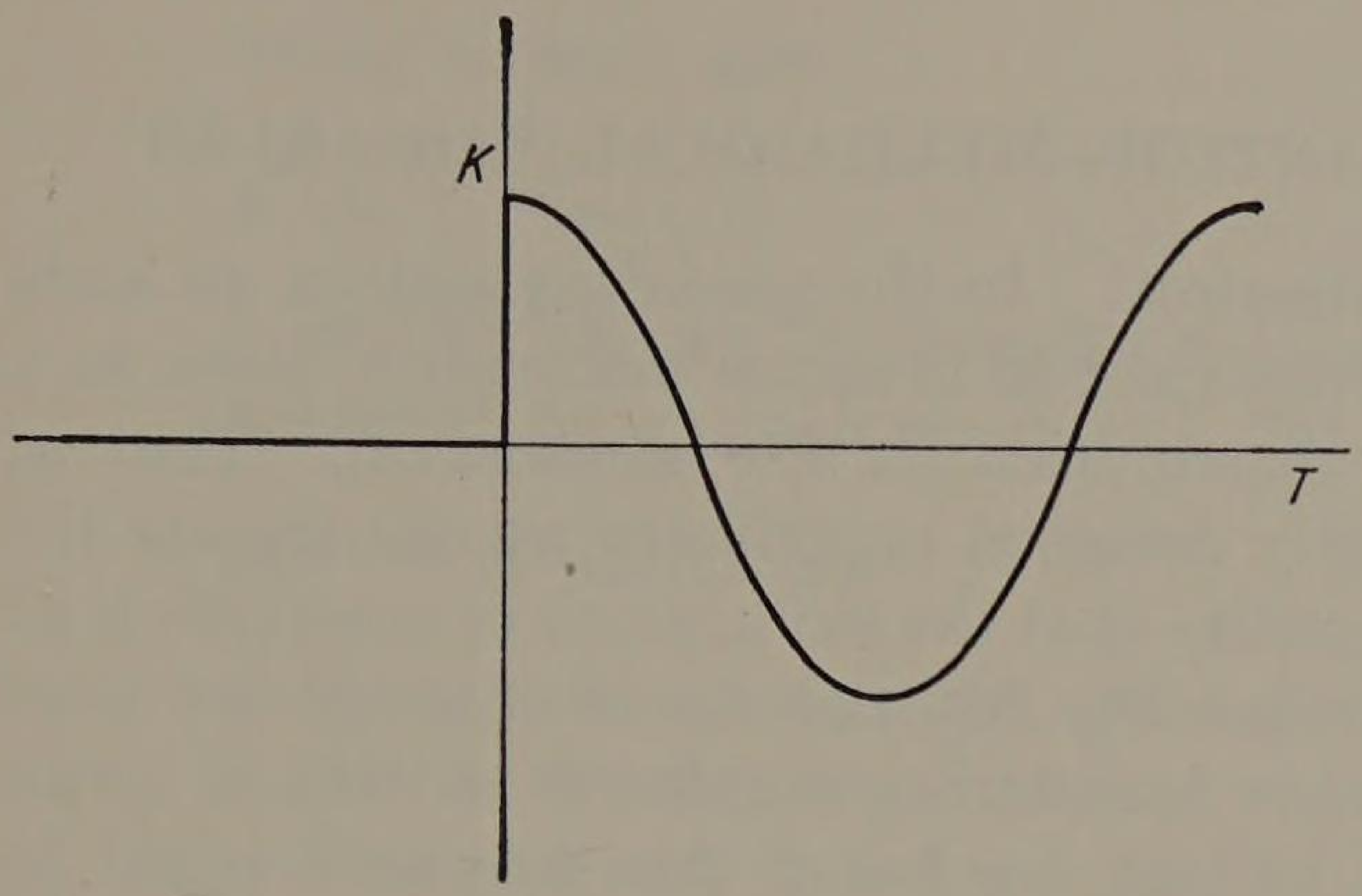


Fig. 5-4 The real part of the free-particle kernel K_0 plotted vs. time. The function is zero for negative times, then starts with a sharp jump at $t = 0$ and continues as a cosine wave of constant amplitude and frequency.

$\mathcal{K}$ is zero for T less than zero, and it suddenly begins to oscillate at $T = 0$. The transformation from time to energy representation is equivalent to a Fourier transformation. Since the wave has a sharp beginning (at $T = 0$), the Fourier transform contains components at all frequencies and thus at all energies. If the function extends over a long time interval (many periods), then one frequency begins to dominate in the Fourier transform. For the free particle this dominating frequency corresponds to the energy $E_0 = p^2/2m$.

It is for this reason that the free-particle kernel contains the factor

$$\frac{i}{E_1{}^2 - p_1{}^2/2m + i\epsilon} = \text{P.P.}\left(\frac{i}{E_1{}^2 - p^2/2m}\right) + \pi\delta\left(E_1{}^2 - \frac{p^2}{2m}\right) \tag{5-19}$$

Here the first term on the right accounts for the transient effects that result from the sudden start at $t = 0$. The second term gives the steady-state behavior and shows that, if we wait long enough, the only energy found is the usual $p^2/2m$; but near $t = 0$ the energy is not given by this classical formula.

Problem 5-2 If we transform only the time and not the spatial variables, defining

$$k(x_2,E_2;x_1,E_1) = \int\int e^{(i/\hbar)E_2t_2}K(x_2,t_2;x_1,t_1)e^{-(i/\hbar)E_1t_1}\,dt_2\,dt_1 \tag{5-20}$$

show that for a system with a time-independent hamiltonian H

$$k(x_2,E_2;x_1,E_1) = 2\pi\hbar i\delta(E_2 - E_1)\sum_m \frac{\phi_m(x_2)\phi_m^*(x_1)}{E_1 - E_m + i\epsilon} \tag{5-21}$$

where ϕ_m and E_m are the eigenfunctions and eigenvalues of H.

5-2 MEASUREMENT OF QUANTUM-MECHANICAL VARIABLES

The Characteristic Function. In the preceding section we have shown how an experiment designed to measure momentum leads to a definition of a probability distribution for the momentum. That is, from the results of a correctly designed experiment we can answer the question: What is the probability that the momentum of a particle is $\mathbf{p}$? From the existence of a probability function for momentum we were led to the discovery of a wave function or amplitude written in terms of momentum variables. In fact, we found that a system could be completely described and problems completely analyzed in a momentum-energy representation as well as the space-time representation which we have used heretofore.

These same results apply to physical variables other than momentum. If any physical quantity can be measured experimentally, a probability function can be associated with it. That is, if an experiment is capable of measuring some characteristic A associated with a system (e.g., the x component of momentum), then after repeating the experiment several times it will be possible to construct the probability function $P(a)$ which gives the probability that in any particular experiment the numerical values of A will be found to be equal to a.

In general, it is possible to associate a probability amplitude with such a probability function. This amplitude would be defined in terms of the measured variable, together with other variables necessary to complete the specification. Let us see what is involved by generalizing our example of a momentum measurement. First we shall take just one dimension, but the extension to several dimensions will be obvious. We ask: Does the system have the property G? For example, G might stand for the statement: The value of the quantity A is equal to a. We must have some way to answer this experimentally. So let us imagine some equipment can be designed so that, if it has the property G, the particle will pass through the equipment and arrive at a certain location on some screen or meter.

The probability of this may be written

$$P(G) = \left|\int K_{exp}(\zeta,x)f(x)\,dx\right|^2 \tag{5-22}$$

if $f(x)$ is the wave function of the system to be measured, $K_{exp}(\zeta,x)$ is the kernel for going through the particular experimental apparatus, and ζ is the position of arrival for particles with the property G. This probability has the alternative mathematical form

$$P(G) = \left|\int g^*(x)f(x)\,dx\right|^2 \tag{5-23}$$

where we have put

$$g^*(x) = K_{exp}(\zeta,x) \tag{5-24}$$

(Defining this as the complex conjugate of a function is just for convenience, as we shall see later.) So we can say

$$\psi(G) = \int g^*(x)f(x)\,dx \tag{5-25}$$

is the amplitude that the system has the property G. This concept is further described in Fig. 5-5.

The property is defined by the function $g^*(x)$ for the following reason. Suppose some other experiment with different equipment, and hence a different kernel $K_{exp'}(\eta,x)$, should be built to measure the same property. In this second experiment the particle arrives at η. Then the probability of finding that the system has the property G is

$$\left|\int K_{exp'}(\eta,x)f(x)\,dx\right|^2 \qquad \text{or} \qquad \left|\int g'^*(x)f(x)\,dx\right|^2 \tag{5-26}$$

Since the property measured is the same, we must obtain the same result in every case for $P(G)$ as we did with the previous experiment. That is to say, we must have

$$\left|\int g'^*(x)f(x)\,dx\right|^2 = \left|\int g^*(x)f(x)\,dx\right|^2 \tag{5-27}$$

for any arbitrary function $f(x)$. This means $g^*(x) = g'^*(x)$ within at least an unimportant constant phase factor $e^{i\delta}$. That is, all methods

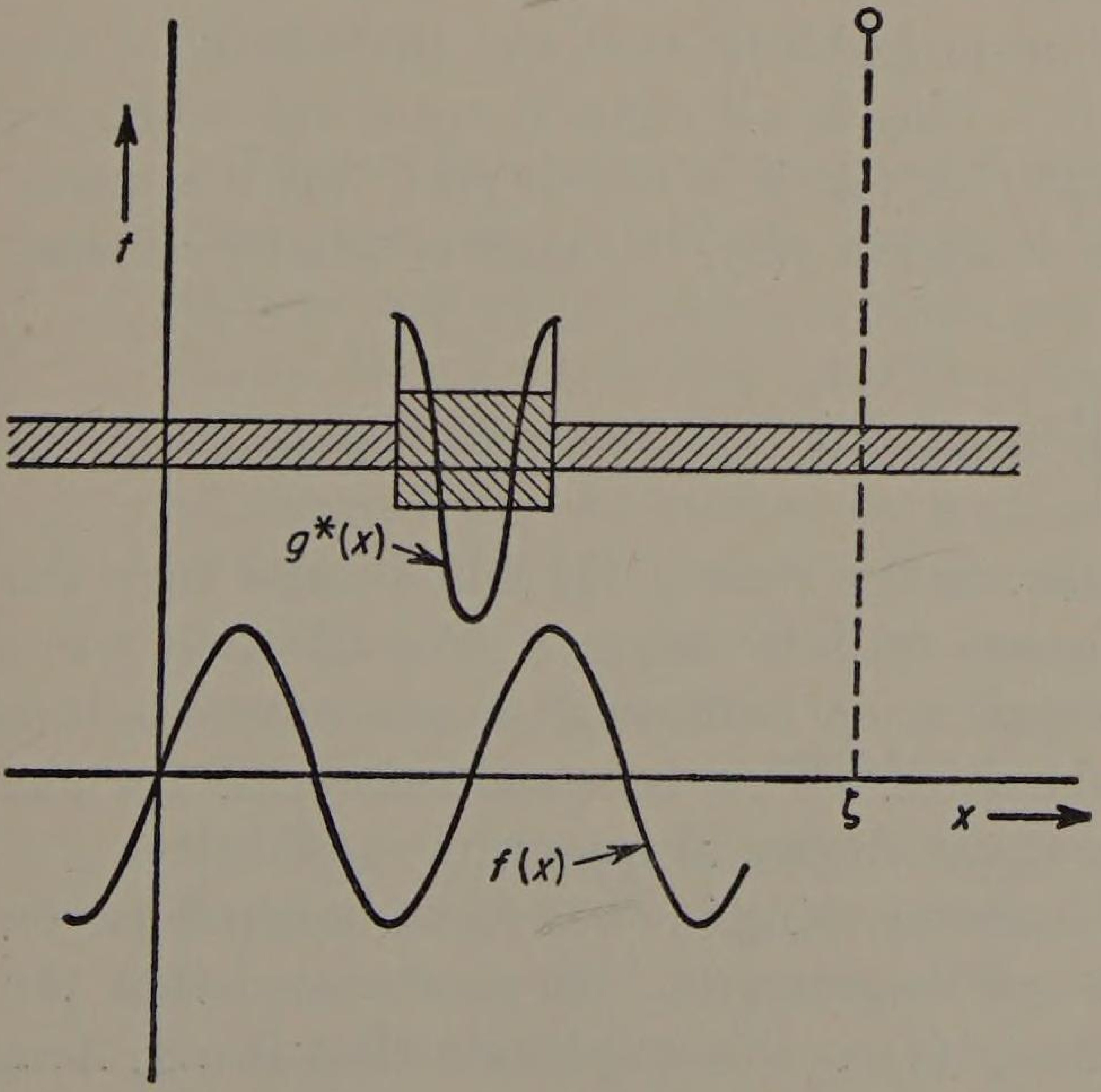


Fig. 5-5 A device designed to measure the property G is placed between the initial point of the incoming particle [with wave function $f(x)$] and the final point $x = \zeta$. The equipment modifies the kernel for the motion (compare Figs. 5-1 and 5-2), making it equal to $g(x)$. The product $f(x)g(x)$, integrated over x, is the amplitude to arrive at ζ after passing through the equipment.

to determine the same property correspond (within a phase) to the same $g^*(x)$. For this reason we call $g^*(x)$ the *characteristic function* of the property G.

We may ask another question. What must the state $f(x)$ be so that it is sure to have the property G? (For example, what is the wave function for a particle whose momentum is definite?) That is, we wish to find an $f(x)$, say, $F(x)$, so that the particle going through the apparatus will certainly arrive at ζ and at no other point ζ'. The amplitude to arrive at ζ' should be proportional to $\delta(\zeta - \zeta')$ (that is, zero unless $\zeta' = \zeta$). Hence

$$\int K_{exp}(\zeta,x)F(x)\,dx = \delta(\zeta - \zeta') \tag{5-28}$$

This we can solve by the relation of the complex conjugate of a kernel to its inverse, discussed in Sec. 4-1. We have from Eq. (4-37)

$$\int K_{exp}(\zeta',x)K^*_{exp}(\zeta,x)\,dx = \delta(\zeta - \zeta') \tag{5-29}$$

so that

$$F(x) = K^*_{exp}(\zeta,x) = g(x) \tag{5-30}$$

That is, $g(x)$ is the wave function of a particle having the property G with certainty. We can say either (1) the particle has the property G or (2) the particle is in the state $g(x)$. So we find: If a particle is in a state $f(x)$, the amplitude it will be found in a state $g(x)$ is

$$\psi(G) = \int g^*(x)f(x)\,dx = \Phi[g(x)] \tag{5-31}$$

For more dimensions, x becomes a space of several variables.

We might say loosely: The probability that the particle is in the state $g(x)$ is $|\int g^*(x)f(x)\,dx|^2$. This is all right if we know what we mean. The system is in state $f(x)$, so it is not in $g(x)$; but if a measurement is made to ask if it is *also* in $g(x)$, the answer will be affirmative with the probability

$$P(G) = \left|\int_{-\infty}^{\infty} g^*(x)f(x)\,dx\right|^2 = \mathcal{P}[g(x)] \tag{5-32}$$

A measurement which asks: Is the state $g(x)$? will always have the answer yes if the wave function actually is $g(x)$. For all other wave functions, repetition of the experiment will result in yes some fraction P (between 0 and 1) of the tries. This is a central result for the probabilistic interpretation of the theory of quantum mechanics.

For all of this we deduce an interesting inverse relationship between a wave function and its complex conjugate. In accordance with the interpretation of Eq. (5-25), $g^*(x)$ is the amplitude that if a system is in position x, then it has the property G. [Such a statement is put mathematically by substituting a δ function for $f(x)$ in Eq. (5-31).] On the other hand, $g(x)$ is the amplitude that if the system has the property G, it is in position x. (This is just a way of giving the definition of a wave function.) One function gives the amplitude for: If A, then B. The other function gives the amplitude for: If B, then A. The inversion is accomplished simply by taking the complex conjugate.

Equation (5-31) can be interpreted as follows: The amplitude that a system has property G is (1) the amplitude $f(x)$ that it is at x times (2) the amplitude $g^*(x)$ that if it is at x, it has property G, with this product summed over the alternatives x.

Problem 5-3 Assume $\int_{-\infty}^{\infty} f^*(x)f(x)\,dx$, which is the probability that a particle of wave function $f(x)$ is somewhere, has been normalized to the value 1. Under this constraint, show that the state $f(x)$ which has the highest probability of having the property G is $f(x) = g(x)$.

Problem 5-4 Suppose the wave function for a system is $\psi(x_1)$ at time t_1. Suppose further that the behavior of the system is described by the kernel $K(x_2,t_2;x_1,t_1)$ for motions in the interval $t_2 \geq t \geq t_1$. Show that the probability that the system is found to be in the state $\chi(x)$ at time t_2 is given by the square of the integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^*(x_2)K(x_2,t_2;x_1,t_1)\psi(x_1)\,dx_1\,dx_2$$

We call this integral the *transition amplitude* to go from state $\psi(x)$ to state $\chi(x)$.

Measurements of Several Variables. In the considerations of the preceding section we assumed an ideal experiment, which means that no quantity besides A could be measured at the same time. That is, we do not allow that more than one $g(x)$ would give the same result, but assert that the maximum possible amount of information has been obtained from the system by a measurement of A.

Now in reality it is common for several variables to determine the state of a system. For example, if only the x component of momentum is measured in a three-dimensional system, no definite $g(x)$ can be defined. Both the wave functions $\exp(ip_x x/\hbar)$ and $\exp(ip_x x/\hbar - ip_y y/\hbar)$ give the same value p_x for the x component of momentum. So if only p_x is measured in a three-dimensional system, the particle could be moving with any component of momentum in the y direction and not change the outcome of the measurement. Nor need the particle come to some unique point in the measuring apparatus. All the particles which arrived at some line or set of points could have the same value for p_x.

Thus in general, we see that the wave function $g(x)$ defines the property G as follows: A state described by the wave function $g(x)$ is certain to have the property G. However, the converse is not necessarily true. That is, it is not certain that all states having the property G are described by the wave function $g(x)$. Only if G includes a specification of all the quantities that may be simultaneously measured is the wave function completely defined by G. Even then there remains an undefined (and unimportant) constant phase factor $e^{i\delta}$.

It is easy to make the necessary extension of the characteristic function $g^*(x)$ when the ideal experiment requires the measurement of more than one variable. Thus suppose we have a set of quantities which we shall call $A, B, C, \ldots$, and which can all be simultaneously measured in an experiment, for example, the x component of momentum, the y component of momentum, etc. Suppose we can completely describe the state of a system by specifying the numerical values $a, b, c, \ldots$ assigned to these quantities. That is, we completely describe the state by saying whether or not it has a certain property. In this case the property in question is that the value of A is a, the value of B is b, etc. Furthermore, suppose that no additional information (information not derivable from a knowledge of the numerical values of A, B, etc.) could be obtained simultaneously by any means.

Imagine we have an experimental setup capable of measuring all these quantities, i.e., capable of telling us whether or not the state has the property that the value of A is a, etc. We shall call the characteristic function of such a property

$$g^*(x) = \chi^*_{a,b,c,\ldots}(x) \tag{5-33}$$

This function is, of course, a function of the numerical values a, b, $c, \ldots$ which the experiment is set up to measure, as well as the coordinate variable x.

Suppose the system is in the state $f(x)$. Then the probability that the experiment would show that the value of A is a, the value of B is b, etc. (i.e., the probability that the state has the property in question), is

$$P(a,b,c,\ldots) = \left|\int \chi^*_{a,b,c,\ldots}(x) f(x)\,dx\right|^2 \tag{5-34}$$

Transformation Functions. Suppose the system is actually in the state $\chi_{a',b',c',\ldots}(x)$, that is, the value of A is a', etc. Then with our experiment the probability of finding the system in a state described by $a, b, c, \ldots$ is zero unless $a = a', b = b', c = c', \ldots$. This means that, with suitable normalizing factors, we have

$$\int_{-\infty}^{\infty} \chi^*_{a,b,c,\ldots}(x)\chi_{a',b',c',\ldots}(x)\,dx = \delta(a-a')\delta(b-b')\delta(c-c') \qquad (5\text{-}35)$$

The function $\chi_{a,b,c,\ldots}(x)$ is the amplitude that if the system is in the state described by $a, b, c, \ldots$, then it will be found at x. The function $\chi^*_{a,b,c,\ldots}(x)$, which we have called the characteristic function, is the amplitude that, if the system is at x, it will be found in the state specified by $a, b, c, \ldots$.

If the system is known to be in the state $f(x)$, then

$$F_{a,b,c,\ldots} = \int_{-\infty}^{\infty} \chi^*_{a,b,c,\ldots}(x)f(x)\,dx \qquad (5\text{-}36)$$

is the amplitude to find the system in the state specified by A having the value a, B having the value b, etc.

The quantities $F_{a,b,c,\ldots}$ are just as good a representation of the state as the function $f(x,y,z,\ldots)$. In fact, if we know the function $F_{a,b,c,\ldots}$, we can reproduce the function $f(x,y,z,\ldots)$ by means of an inverse transformation.

The function $F_{a,b,c,\ldots}$ is called the $A, B, C, \ldots$ representation of the state. (In the preceding section we had an example of this in the momentum representation.) The function $f(x,y,z,\ldots)$ is the customary coordinate representation, or $x, y, z, \ldots$ representation, of the state. Transformations between the two are carried out with the help of the functions χ and χ^*. In particular, the function $\chi^*_{a,b,c,\ldots}(x,y,z,\ldots)$ is the transformation function necessary to go from the $x, y, z, \ldots$ representation to the $A, B, C, \ldots$ representation, while the function $\chi_{a,b,c,\ldots}(x,y,z,\ldots)$ is the transformation function necessary to go the other direction. Thus the inverse of the transformation given by Eq. (5-36) is

$$f(x,y,z,\ldots) = \sum_a\sum_b\sum_c \cdots F_{a,b,c,\ldots}\chi_{a,b,c,\ldots}(x,y,z,\ldots) \qquad (5\text{-}37)$$

This says that the amplitude to be found at x is the amplitude $F_{a,b,c,\ldots}$ to be found with $A = a, B = b, \ldots$ times the amplitude $\chi_{a,b,c,\ldots}(x)$ to be at x if $A = a, B = b$, etc., summed over alternatives $a, b, c, \ldots$.

Problem 5-5 Assume that the function $f(x,y,z,\ldots)$ can be represented by

$$f(x,y,z,\ldots) = \sum_a\sum_b\sum_c \cdots F'_{a,b,c,\ldots}\chi_{a,b,c,\ldots}(x,y,z,\ldots) \qquad (5\text{-}38)$$

By substituting this relation into Eq. (5-36), and using the orthogonal properties of χ as defined by Eq. (5-35), show that $F'_{a,b,c,\ldots} = F_{a,b,c,\ldots}$.

Problem 5-6 Suppose A, B, and C are the three cartesian components of momentum p_x, p_y, p_z. What is the form of the function $\chi_{a,b,c}(x,y,z)$? Using the results of Sec. 5-2, verify the relations obtained in Sec. 5-1.

Problem 5-7 Suppose that the $A, B, C, \ldots$ representation does not correspond to either coordinate representation or momentum representation, but instead is some third way of representing the state of the system. Suppose we know the function $\chi_{a,b,c,\ldots}(x,y,z,\ldots)$ which permits us to transform back and forth between coordinate representation and $A, B, C, \ldots$ representation. Suppose further that we know the transformation function necessary to transform back and forth between coordinate representation and momentum representation. What then is the function necessary for the transformation between momentum representation and $A, B, C, \ldots$ representation?

5-3 OPERATORS

Expected Values. We can develop a few further properties of these transformation functions. Let us try to answer this question. A system is in a state specified by the wave function $f(x)$, and the quantity A is measured. If the measurement is repeated many times, what is the average value which will be obtained for A? We shall denote this average value (sometimes called *expected value*) by the symbol $\langle A \rangle$.

Suppose it is possible, in principle, to measure simultaneously several physical quantities $A, B, C, \ldots$, where a measurement of A could produce any one of a continuous or discrete set of values a, a measurement of B could produce any one of a continuous or discrete set of values b, etc. The probability of obtaining one particular set of values $a, b, c, \ldots$ is $|F_{a,b,c,\ldots}|^2$. So the probability of obtaining a particular value a in a measurement of A, irrespective of the values taken on by $B, C, \ldots$ (for example, if $B, C, \ldots$ were not measured at all) is

$$P(a) = \sum_b \sum_c \cdots |F_{a,b,c,\ldots}|^2 \tag{5-39}$$

In this equation summations are carried out over all possible values in the continuous or discrete sets of $b, c, \ldots$.

The average, or expected, value resulting from a measurement of A is obtained by multiplying the probability of Eq. (5-39) by a and summing the result over all possible values of a. Thus

$$\langle A \rangle = \sum_a \sum_b \sum_c \cdots a |F_{a,b,c,\ldots}|^2 \tag{5-40}$$

The need for computing such expected values arises frequently in quantum-mechanical problems. It is useful to have available formulas which simplify such computations. This subject, the subject of operators, was discussed briefly in Sec. 4-1. Now we shall develop a few additional results. However, nowhere in this book shall we attempt a really thorough study of operator calculus, since several excellent works along this line are already available.[1]

The Operator. Let us try to express the expected value of A directly in terms of the original wave function $f(x)$. Note first that the absolute square of $F_{a,b,c,\ldots}$ can be written as

$$|F_{a,b,c,\ldots}|^2 = F^*_{a,b,c,\ldots} F_{a,b,c,\ldots} \tag{5-41}$$

Then, using Eq. (5-36), we can write

$$\begin{aligned} \langle A \rangle &= \sum_a \sum_b \sum_c \cdots a \int_{-\infty}^{\infty} \chi_{a,b,c,\ldots}(x) f^*(x)\, dx \int_{-\infty}^{\infty} \chi^*_{a,b,c,\ldots}(x') f(x')\, dx' \\ &= \int_{-\infty}^{\infty} f^*(x) R(x)\, dx \end{aligned} \tag{5-42}$$

In the second line of this equation we have made use of the substitution

$$R(x) = \int_{-\infty}^{\infty} G_A(x,x') f(x')\, dx' \tag{5-43}$$

where we have written

$$G_A(x,x') = \sum_a \sum_b \sum_c \cdots a \chi_{a,b,c,\ldots}(x) \chi^*_{a,b,c,\ldots}(x') \tag{5-44}$$

Equation (5-43) says that the function $R(x)$ results from the function $f(x)$ as the result of an integration performed with the help of a linear integral operator $G_A(x,x')$ associated with the quantity A.

---

[1] For example, see Dirac, "Principles of Quantum Mechanics," Clarendon Press, Oxford, 1947.

Often an equation like Eq. (5-43) is symbolized by the notation

$$R = \mathcal{A}f \tag{5-45}$$

where $\mathcal{A}$ stands for a linear operator which operates on the function f. In the present case $\mathcal{A}$ stands for the operation displayed on the right-hand side of Eq. (5-43), that is, multiplication by the function G_A and integration. The operator $\mathcal{A}$ is associated with the physical quantity A. Using this notation, we can write

$$\langle A \rangle = \int_{-\infty}^{\infty} f^*(x)\mathcal{A}f(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^*(x)G_A(x,x')f(x')\,dx'\,dx \tag{5-46}$$

Problem 5-8 Note that Eq. (5-44) implies $G_A^*(x,x') = G_A(x',x)$. With this in mind show that for any two wave functions $g(x)$ and $f(x)$, both of which approach 0 as x goes to ∞,

$$\int_{-\infty}^{\infty} g^*(x)\mathcal{A}f(x)\,dx = \int_{-\infty}^{\infty} [\mathcal{A}g(x)]^*f(x)\,dx \tag{5-47}$$

Any operator, such as $\mathcal{A}$, for which Eq. (5-47) holds is called *hermitian* [cf. Eq. (4-30)].

Problem 5-9 The transformation function between space representation and momentum representation is

$$\chi_{a,b,c}(\mathbf{r}) = e^{(i/\hbar)\,\mathbf{p}\cdot\mathbf{r}} \tag{5-48}$$

(see Prob. 5-6). Choose the physical quantity A as the momentum p_x in the x direction. Show that the function G_A is

$$G_{p_x}(x,x') = \frac{\hbar}{i}\,\delta'(x - x')\delta(y - y')\delta(z - z') \tag{5-49}$$

where $\delta'(x) = \dfrac{d}{dx}\,\delta(x)$. With this result determine the operator corresponding to the x component of momentum and show that the expected value of this component of momentum can be written as

$$\langle p_x \rangle = \int_{-\infty}^{\infty} f^*(x)\,\frac{\hbar}{i}\frac{\partial f}{\partial x}\,dx \tag{5-50}$$

Problem 5-10 Suppose the quantity A corresponds to the x coordinate of position. Show that the correct formulas for the mean value of x result when the function $G_A(x,x')$ is taken to be

$$G_x(x,x') = x\delta(x - x')\delta'(y - y')\delta'(z - z') \tag{5-51}$$

and the operator corresponding to x is simply multiplication by x, that is,

$$\mathcal{X}f(x) = xf(x) \tag{5-52}$$

Eigenfunctions and Eigenvalues. The wave functions $\chi_{a,b,c,\ldots}$, as discussed in Sec. 5-2, show a particularly simple behavior when subjected to the operation $\mathcal{A}$. Thus

$$\mathcal{A}\chi_{a,b,c,\ldots}(x) = a\chi_{a,b,c,\ldots}(x) \tag{5-53}$$

Problem 5-11 Show that this last result is true.

When a function χ satisfies an equation such as (5-53), we say that χ is an eigenfunction of the operator $\mathcal{A}$ associated with the eigenvalue a.

If two physical quantities can be simultaneously measured, then the operators associated with these quantities, $\mathcal{A}$ and $\mathcal{B}$, for example, satisfy an interesting relationship, namely, $\mathcal{A}(\mathcal{B}f) = \mathcal{B}(\mathcal{A}f)$. This relation says that the result of performing one operation after the other is the same regardless of the order in which the operations are performed. In this case the two operators are said to *commute:*

$$\mathcal{A} \cdot \mathcal{B} = \mathcal{B} \cdot \mathcal{A}$$

In general, we cannot expect the commutation relation to hold between operators, but in this special case it does. The reason for this is that if A and B are physical quantities which can be measured simultaneously, they can form part of a set $A, B, C, \ldots$ of simultaneously measurable quantities with a single characteristic function $\chi_{a,b,c,\ldots}$. If the operator $\mathcal{B}$ is substituted for $\mathcal{A}$ and the value b is substituted for a in Eq. (5-53), the result is still valid, so

$$\mathcal{A}(\mathcal{B}\chi) = \mathcal{A}(b\chi) = b(\mathcal{A}\chi) = ba\chi = ab\chi \tag{5-54}$$

which is true, since a, b are just numbers. Now also

$$\mathcal{B}(\mathcal{A}\chi) = \mathcal{B}(a\chi) = a(\mathcal{B}\chi) = ab\chi \tag{5-55}$$

A comparison of these two equations proves the commutation of the operators $\mathcal{A}$ and $\mathcal{B}$ when acting upon any of the functions $\chi_{a,b,c,\ldots}$. Since both these operations are linear (i.e., they do not involve computations with higher powers of the function χ), the commutation relation must apply to any linear combination of the χ functions.

If the χ functions comprise a "complete set" (which is typical) we can construct any function at all from such a linear combination. So the operation AB and the operation BA give the same result on any function; that is, they commute.

Problem 5-12 Show that the x coordinate of position and the x coordinate of momentum are not simultaneously measurable quantities.

There are situations in which a set of commuting mathematical operations $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, . . . are already known and it is required to find the functions (the eigenfunctions) which are associated with them. This requires solving a set of equations such as

$$\mathcal{A}\chi = a\chi \qquad \mathcal{B}\chi = b\chi \qquad \mathcal{C}\chi = c\chi \qquad \cdots \tag{5-56}$$

For example, suppose the operators for momentum in the x, y, z directions p_x, p_y, and p_z are given as $(\hbar/i)\,\frac{\partial}{\partial x}$, $(\hbar/i)\,\frac{\partial}{\partial y}$, $(\hbar/i)\,\frac{\partial}{\partial z}$. What are the eigenfunctions of this set of operators corresponding to a state in which p_x has the value a, p_y has the value b, and p_z has the value c (these are, of course, the eigenvalues)? We must solve the equations

$$-\frac{\hbar}{i}\frac{\partial\chi}{\partial x} = a\chi \qquad -\frac{\hbar}{i}\frac{\partial\chi}{\partial y} = b\chi \qquad -\frac{\hbar}{i}\frac{\partial\chi}{\partial z} = c\chi \tag{5-57}$$

and the solution is some arbitrary constant times $e^{(i/\hbar)(ax+by+cz)}$. This agrees with our previous knowledge that a particle having a known momentum $\mathbf{p}$ has the wave function $e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}}$.

Interpretation of Energy Expansion. Various expressions involving $\phi_n(x)$ can be interpreted more completely now. For example, consider the expansion in Eq. (4-59) of the kernel in terms of the solutions ϕ_n of a constant hamiltonian

$$K(x_2,t_2;x_1,t_1) = \sum_n \phi_n(x_2)\,\phi_n^*(x_1)e^{-(i/\hbar)E_n(t_2-t_1)} \tag{5-58}$$

We notice first that $\phi_n(x)$ is the amplitude that if we are in state n, we are at position x. Therefore, from our previous discussion (cf. Sec. 5-2) $\phi_n^*(x)$ is the amplitude that if we are at x, we are in state n. Now let us interpret Eq. (5-58) this way. The amplitude to get from position 1 at time t_1 to position 2 at time t_2 is the sum over alternatives. This time the alternatives will be divided into the various energy states in which the transition can be made. Thus we must sum over all of the states n the product of the following terms:

1. $\phi_n^*(x_1)$, which is the amplitude that if we are at x_1, then we are in the state n.
2. $e^{-(i/\hbar)E_n(t_2-t_1)}$, which is the amplitude to be in state n at the time t_2 if we are in the state n at the time t_1.†
3. $\phi_n(x_2)$, which is the amplitude to be found at x_2 when we know we are in the state n.

Problem 5-13 Discuss the possibility of interpreting $\phi_n(x)$ as a $\chi_{a,b,c,\ldots}(x)$ function discussed in Sec. 5-2. That is, say $\phi_n(x)$ is the transformation function to go from the x representation to a representation specified by n (energy representation).

---

† There is no amplitude to change the state. That is the importance of these particular functions ϕ_n.

6

The Perturbation Method in Quantum Mechanics

If a quantum-mechanical system is subjected to a potential which introduces only quadratic terms into the action, then we have seen in Chap. 3 how the resulting motion can be determined with the path integral method. However, many of the interesting potentials which arise in quantum-mechanical problems are not of this special type and cannot be handled so easily. In this chapter we shall develop a method of treating more complicated potentials. The method which we discuss, called the *perturbation expansion*, is most useful when the potential is comparatively weak (compared, for instance, to the kinetic energy of the system).

Although the perturbation expansion can be developed along strictly mathematical lines, it is capable of an interesting physical interpretation. This interpretation, which we shall also present, leads to a deeper understanding of quantum-mechanical behavior.

In the second section of this chapter we shall undertake a special application of the perturbation method. We shall consider the motion of an electron when it is scattered by an atom. In describing the scattering interaction we shall find useful the classical notion of a cross-sectional area which the atom presents to the impinging electron. Although this area is related to the actual size of the atom, we shall find that its complete description depends upon the quantum-mechanical aspects of the interacting system.

6-1 THE PERTURBATION EXPANSION

The Terms of the Expansion. Suppose a particle is moving in a potential $V(x,t)$. For the present, the motion will be restricted to one dimension. Then the kernel for motion between the points a and b is

$$K_V(b,a) = \int_a^b \left(\exp \left\{ \frac{i}{\hbar} \int_{t_a}^{t_b} \left[\frac{m}{2} \dot{x}^2 - V(x,t) \right] dt \right\} \right) \mathcal{D}x(t) \tag{6-1}$$

The subscript notation K_V is used to remind us that the particle is in the potential V. From now on the notation K_0 will denote the kernel for the motion of a free particle.

In some cases the kernel K_V can be determined by the methods already studied. For instance, in Chap. 3 we determined the kernel for the harmonic oscillator subject to an outside force $f(t)$. Here the potential was [see Eq. (3-65)]

$$V(x,t) = \frac{m}{2} \omega^2 x^2 - x f(t) \tag{6-2}$$

In general, we have found that if the potential is quadratic in x, the kernel can be determined exactly, whereas if it is sufficiently slowly varying, the semiclassical approximation is adequate. There are some other types of potentials which can be successfully treated with the help of Schrödinger's equation. Now we are studying a technique which is often useful if the effect of the potential is small.

Suppose the potential is small, or more precisely, suppose the time integral of the potential along a path is small compared to $\hbar$. Then the part of the exponential of Eq. (6-1) which depends upon $V(x,t)$ can be expanded as

$$\exp\left[-\frac{i}{\hbar}\int_{t_a}^{t_b} V(x,t)\,dt\right] = 1 - \frac{i}{\hbar}\int_{t_a}^{t_b} V(x,t)\,dt + \frac{1}{2!}\left(\frac{i}{\hbar}\right)^2\left[\int_{t_a}^{t_b} V(x,t)\,dt\right]^2 \cdots \tag{6-3}$$

which is defined along any particular path $x(t)$.

Using this expansion in Eq. (6-1) results in

$$K_V(b,a) = K_0(b,a) + K^{(1)}(b,a) + K^{(2)}(b,a) + \cdots \tag{6-4}$$

where

$$K_0(b,a) = \int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,dt\right)\right]\mathfrak{D}x(t) \tag{6-5}$$

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,dt\right)\right]\int_{t_a}^{t_b} V[x(s),s]\,ds\,\mathfrak{D}x(t) \tag{6-6}$$

$$K^{(2)}(b,a) = -\frac{1}{2\hbar^2}\int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,dt\right)\right]\int_{t_a}^{t_b} V[x(s)]\,ds \times \int_{t_a}^{t_b} V[x(s'),s']\,ds'\,\mathfrak{D}x(t) \tag{6-7}$$

and so forth. To avoid confusion in the integrals over V, we call the time variables s, s', etc.

Evaluation of the Terms. First consider the kernel $K^{(1)}$. We wish to interchange the order of integration over the variable x and the path $x(t)$. We write

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int_{t_a}^{t_b} F(s)\,ds \tag{6-8}$$

where

$$F(s) = \int_a^b \left[\exp\left(\frac{i}{\hbar}\int_{t_a}^{t_b}\frac{m\dot{x}^2}{2}\,dt\right)\right] V[x(s),s]\,\mathfrak{D}x(t) \tag{6-9}$$

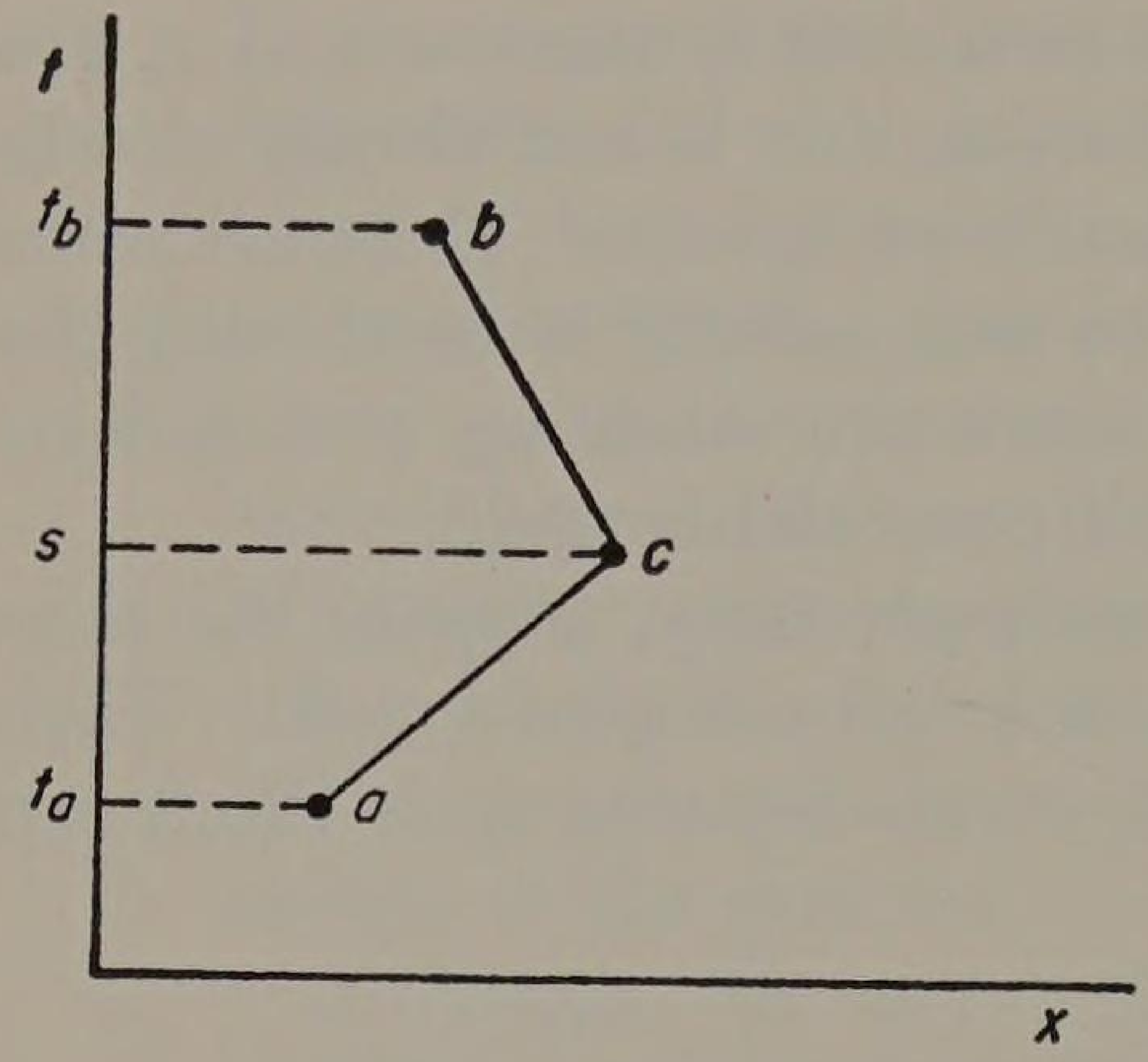


Fig. 6-1 A particle starts from a and moves as a free particle to c. Here it is acted upon, or scattered, by the potential $V[x(s),s] = V_c$. Thereafter it moves as a free particle to b. The amplitude for such a motion is given in Eq. (6-10). If this amplitude is integrated over all possible positions of the point c, the result is the first-order term in the perturbation expansion.

The path integral $F(s)$ can be described as follows. It is the sum over all paths of the free-particle amplitude. However, each path is weighted by the potential $V[x(s),s]$ evaluated at the time s. The only characteristic of the path $x(t)$ which is involved in this particular V is the position of the path at the particular time $t = s$. This means that before and after the time s the paths involved in $F(s)$ are the paths of an ordinary free particle. The situation is sketched in Fig. 6-1.

Using the same arguments which led to Eq. (2-31), we divide each path into two parts, one before the time $t = s$ and one after this time. To be specific, we shall assume that each path goes through the point x_c at this division time. Later on we shall integrate over all values of x_c. If we denote the point $x_c(s)$ by c (that is, $s = t_c$), then the sum over all such paths can be written as $K_0(b,c)K_0(c,a)$. This means that $F(s) = F(t_c)$ can be written as

$$F(t_c) = \int_{-\infty}^{\infty} K_0(b,c)V(x_c,t_c)K_0(c,a)\,dx_c \tag{6-10}$$

Substituting this into Eq. (6-8) gives [with $V(c) = V(x_c,t_c)$]

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int_{t_a}^{t_b}\int_{-\infty}^{\infty} K_0(b,c)V(c)K_0(c,a)\,dx_c\,dt_c \tag{6-11}$$

Here the limits on the integral over x have been written as $\pm\infty$. In a practical problem the limits will be established by the potential (which in most cases drops to 0 when x becomes very large) or by the equipment, which restricts the range of x.

Interpretation of the Terms. Equation (6-11) is very important and very useful, so we shall develop a special interpretation to help think about it physically. We call the interaction between the potential and the particle a *scattering;* thus we say that the potential scatters the particle and that the *amplitude to be scattered by a potential is* $-(i/\hbar)V$ *per unit volume and per unit time.*

With this interpretation we can describe K_V in the following way. K_V is, of course, a sum over alternative ways in which the particle may move from point a to point b. The alternatives are:

1. The particle may not be scattered at all $[K_0(b,a)]$.
2. The particle may be scattered once $[K^{(1)}(b,a)]$.
3. The particle may be scattered twice $[K^{(2)}(b,a)]$.

   Etc.

In accordance with this interpretation, the various paths of the particles are diagramed in Fig. 6-2.

Each one of these alternatives is itself a sum over alternatives. Consider, for example, the kernel for a single scattering, $K^{(1)}(b,a)$.

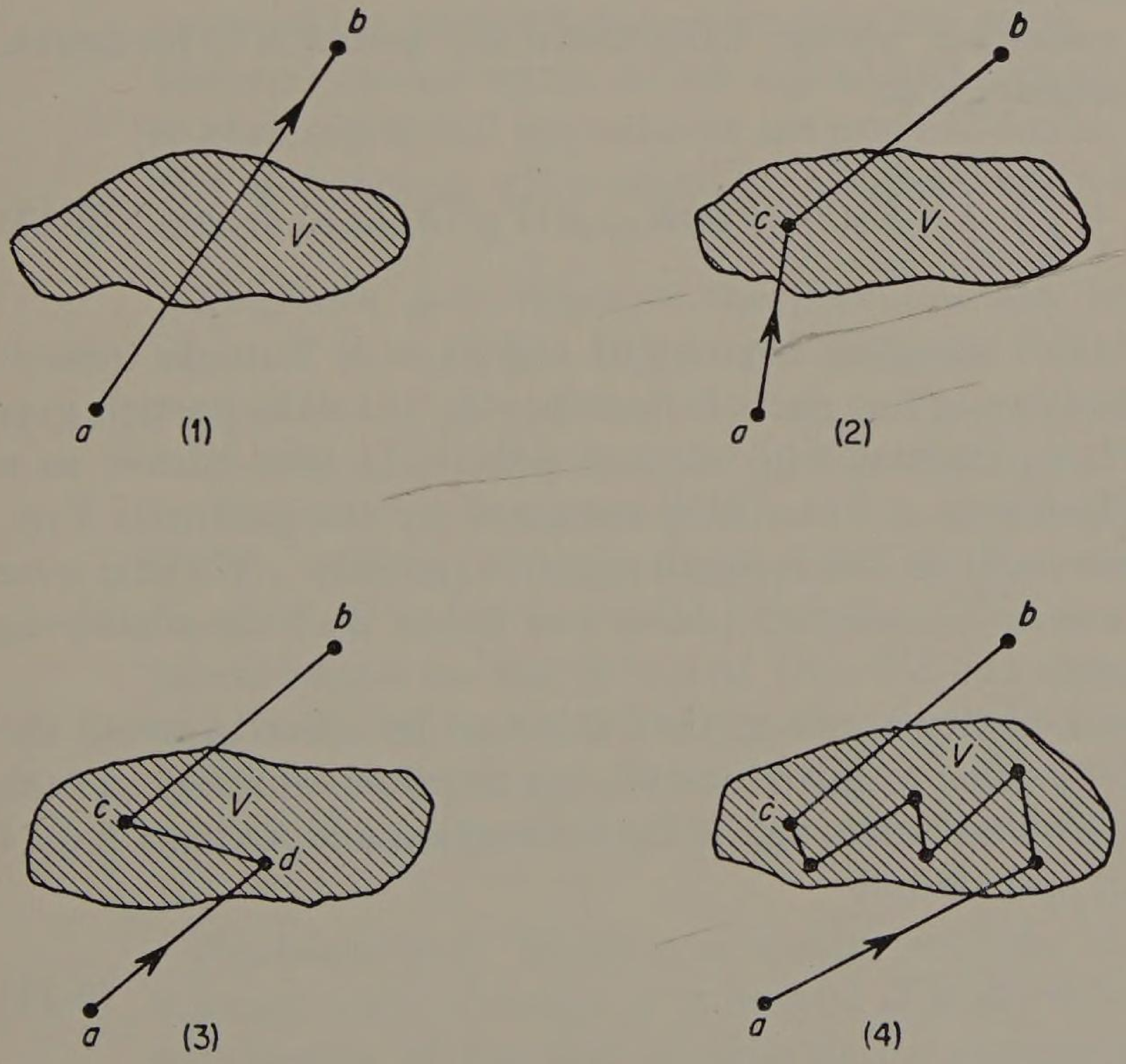


Fig. 6-2 In (1) a particle moves from a to b through the potential V without being scattered. The amplitude for this is $K_0(b,a)$. In (2) the particle is scattered once at c as it moves through the potential V. The amplitude for this is $K^{(1)}(b,a)$. In (3) the particle is scattered twice with the amplitude $K^{(2)}(b,c)$, and in (4) it is scattered n times, the last scattering taking place at c. The total amplitude for motion from a to b with any number of scatterings is $K_0 + K^{(1)} + K^{(2)} + \cdots + K^{(n)} + \cdots$.

One of the alternatives which comprise this kernel consists of the following motion. The particle starts from a, moves as a free particle to the point x_c, $t_c = c$, is there scattered by the potential $V(c)$, after which it moves as a free particle from the point c to the final position b. The amplitude for such a path is

$$K_0(b,c)\left[-\frac{i}{\hbar}V(c)\,dx_c\,dt_c\right]K_0(c,a) \tag{6-12}$$

(It must be remembered that in the convention we are using the motion of the particle is traced by reading the formulas backward, i.e., from right to left.)

The construction of this amplitude follows the rule stated in Sec. 2-5, namely, that the amplitudes for events occurring in succession in time multiply. The completed form for the kernel $K^{(1)}$ is obtained by adding up all such alternatives by integrating over x_c and t_c, as indicated in Eq. (6-11).

Using this reasoning, we can write down the kernel $K^{(2)}$ for double scattering immediately as

$$K^{(2)}(b,a) = \left(-\frac{i}{\hbar}\right)^2 \iint K_0(b,c)V(c)K_0(c,d)V(d)K_0(d,a)\,d\tau_c\,d\tau_d \tag{6-13}$$

where $d\tau = dx\,dt$. Reading from right to left, this formula means: The particle moves as a free particle from a to d. At d the particle gets scattered by the potential $V(d)$ at that point. It then moves as a free particle from d to c, where it is scattered by the potential $V(c)$. After that it moves from c to b, again as a free particle. We sum over all the alternatives, namely, all places and times that the scattering may take place.

Here we have tacitly assumed that $t_c > t_d$. In order to avoid the complication of having to introduce this assumption explicitly in each such example, we shall make use of the convention adopted in Chap. 4 [cf. Eq. (4-28)] and assume

$$K(b,a) = 0 \qquad \text{for } t_b < t_a \tag{6-14}$$

Then Eq. (6-13) is correct without any restrictions on the range of integration of t_c and t_d.

The reader may wonder what happened to the factor ½ which appears in Eq. (6-7) but is apparently omitted in Eq. (6-13). Note that in Eq. (6-13) the range of integration for t_d is still from t_a to t_b;

however, the range of t_c has been restricted [by the definition of Eq. (6-14)] to lie between t_a and t_b. This restriction cuts the value of the double integral in half. To see this more clearly, suppose the double integral of Eq. (6-7) is rewritten as

$$\int_{t_a}^{t_b}\int_{t_a}^{t_b} V[x(s),s]V[x(s'),s']\,ds'\,ds = \int_{t_a}^{t_b}\int_{s}^{t_b} V[x(s),s]V[x(s'),s']\,ds'\,ds + \int_{t_a}^{t_b}\int_{t_a}^{s} V[x(s),s]V[x(s'),s']\,ds'\,ds \tag{6-15}$$

The first term on the right-hand side of this equation satisfies the restrictions implied by Eq. (6-14). By interchanging the order of integration, the second term on the right-hand side can be rewritten as

$$\int_{t_a}^{t_b}\int_{s'}^{t_b} V[x(s),s]V[x(s'),s']\,ds\,ds' \tag{6-16}$$

If the variables s and s' are interchanged in the last expression, the value of the double integral remains the same. This means that the first and second terms on the right-hand side of Eq. (6-15) are equal, so each one is half the value of the original double integral. This same sort of argument will account for a factor $1/n!$ in an expression for $K^{(n)}$.

Problem 6-1 Suppose the potential can be written as $U + V$, where V is small but U is large. Suppose further that the kernel for motion in the potential of U alone can be worked out (for example, U might be quadratic in x and independent of time). Show that the motion in complete potential $U + V$ is described by Eqs. (6-4), (6-11), (6-13), and (6-14) with K_0 replaced by K_U, where K_U is the kernel for motion in the potential U alone. Thus we can consider V as a perturbation on the potential U. We can say that $-(i/\hbar)V$ is the amplitude to be scattered by the perturbed part of the potential (per unit volume and per unit time). K_U is the amplitude for the motion in the system in the unperturbed potential U.

Problem 6-2 Suppose a system consists of two particles which interact only through a potential $V(x,y)$, where x represents the coordinates of the first particle and y represents the coordinates of the second [cf. Sec. 3-8 and Eq. (3-75)]. Apart from this interaction, the particles are free. If V were 0, then K would be simply a product of the two free-particle kernels. Using this fact, develop a perturbation expansion for $K_V(x_b,y_b,t_b;x_a,y_a,t_a)$. By what rules of physical reasoning can the various terms in this expansion be described?

6-2 AN INTEGRAL EQUATION FOR K_V

Before applying the results of the preceding paragraphs to a special example, we shall develop some mathematical relations involving the kernels and wave functions of systems moving in a potential field. Using the results so far obtained, we can write Eq. (6-4) as follows:

$$K_V(b,a) = K_0(b,a) - \frac{i}{\hbar}\int K_0(b,c)V(c)K_0(c,a)\,d\tau_c$$
$$+ \left(-\frac{i}{\hbar}\right)^2 \iint K_0(b,c)V(c)K_0(c,d)V(d)K_0(d,a)\,d\tau_c\,d\tau_d + \cdots \tag{6-17}$$

Alternatively, this expression could be written as

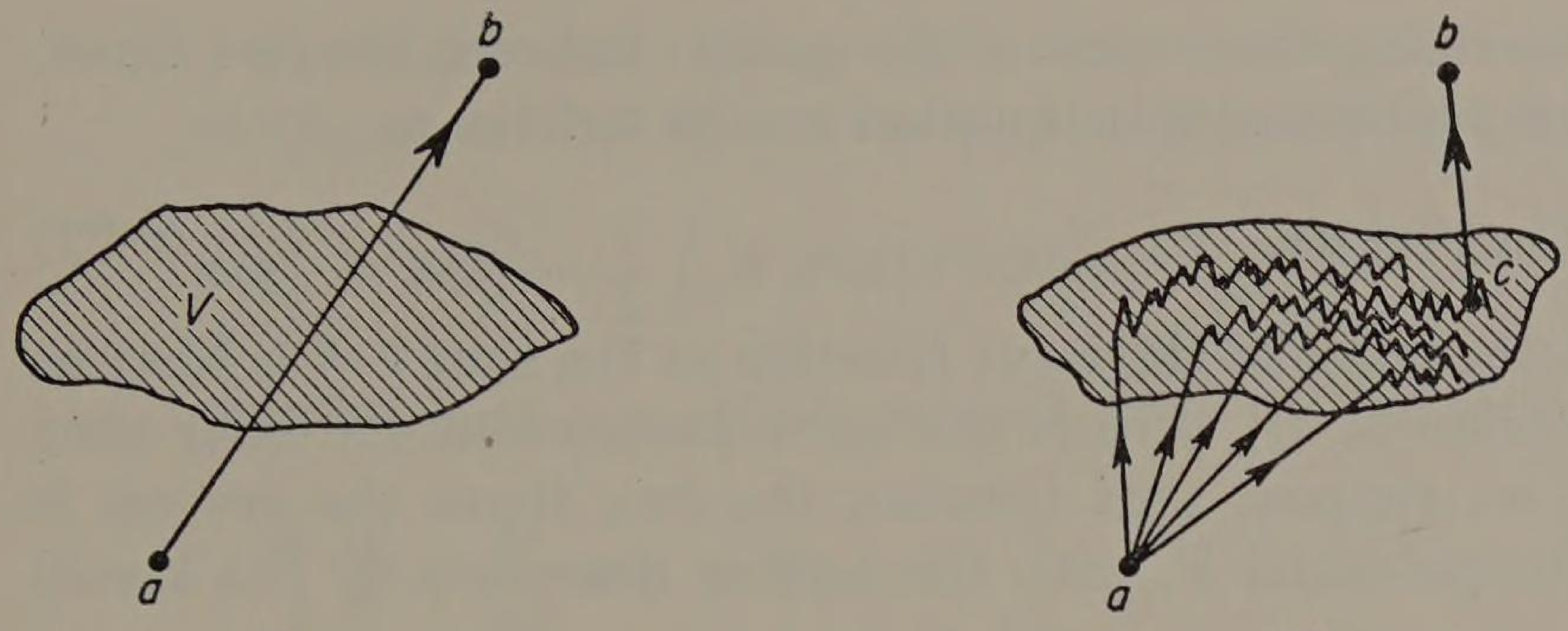
$$K_V(b,a) = K_0(b,a) - \frac{i}{\hbar}\int K_0(b,c)V(c)\left[K_0(c,a) - \frac{i}{\hbar}\int K_0(c,d)V(d)K_0(d,a)\,d\tau_d + \cdots\right]d\tau_c \tag{6-18}$$

The expression in the brackets has the same form as Eq. (6-17). In both cases the sums extend over an infinite number of terms. This means that K_V can be written as

$$K_V(b,a) = K_0(b,a) - \frac{i}{\hbar}\int K_0(b,c)V(c)K_V(c,a)\,d\tau_c \tag{6-19}$$

which is an exact expression. This is an integral equation determining K_V if K_0 is known. (Note that for the situation described in Prob. 6-1, K_0 would be replaced by K_U.) Thus the path integral problem has been transformed into an integral equation.

This last result can be understood physically in the following way. The total amplitude for the transition of the system from a to b, with any number of scatterings, can be expressed as the sum of two alternatives. The first alternative is the amplitude that the transition takes place with no scatterings, which is expressed by K_0. The second alternative is the amplitude that the transition takes place with one or more scatterings, which is given by the last term of Eq. (6-19). In this last term the point c can be thought of as the point at which the *last* scattering takes place. Thus the system moves from a to c in the potential field with its motion exactly described by $K_V(c,a)$. Then at the point c the final scattering takes place, after which the system moves as a free system (without scattering) to the



(1) (2)

Fig. 6-3 In (1) the particle moves from a to b through the potential V as a free particle, described by the amplitude $K_0(b,a)$. In (2) the particle is scattered one or more times by V, with the last scattering taking place at c. The motion from a to c is described by $K_V(c,a)$, and that from c to b by $K_0(b,c)$. A combination of the two situations, when all positions of c are accounted for, covers all possible cases and gives $K_V(b,a)$ in the form of Eq. (6-19).

point b, as represented by the kernel K_0. This interpretation is diagramed in Fig. 6-3.

Since the last scattering could take place at any point in space and time between a and b, the amplitude for this composite motion, represented by the integrand of the last term of Eq. (6-19), must be integrated over all possible positions of the point c.

Problem 6-3 For a free particle, Eq. (4-29) reduces to

$$-\frac{\hbar}{i}\frac{\partial}{\partial t_b}K_0(b,a)+\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_b^2}K_0(b,a)=i\hbar\delta(t_b-t_a)\delta(x_b-x_a) \tag{6-20}$$

Show, from this result and Eq. (6-19), that the kernel K_V satisfies the differential equation

$$-\frac{\hbar}{i}\frac{\partial}{\partial t_b}K_V(b,a)+\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_b^2}K_V(b,a)+V(b)K_V(b,a) = i\hbar\delta(x_b-x_a)\delta(t_b-t_a) \tag{6-21}$$

6-3 AN EXPANSION FOR THE WAVE FUNCTION

In Sec. 3-4 we introduced the idea of a wave function and discussed some relations between wave functions and kernels. Equation (3-42) of that section shows how the wave function at the time t_b can be obtained from the wave function at an earlier time t_a with the help of the kernel describing the motion of the system between the two times. For our present purposes this equation can be written as

$$\psi(b) = \int K_V(b,a)f(a)\,dx_a \tag{6-22}$$

where $f(a)$ is the value of the wave function at the time $t = t_a$ [that is, $f(a)$ is a function of x_a], $\psi(b)$ is the wave function at the later time $t = t_b$,† and we suppose that between the two times the system is moving in the potential V, with the motion described by the kernel $K_V(b,a)$.

If the series expansion [Eq. (6-18)] for K_V is substituted into this equation, the result will be a series expansion for $\psi(b)$. Thus

$$\psi(b) = \int K_0(b,a)f(a)\,dx_a - \frac{i}{\hbar}\int\int K_0(b,c)V(c)K_0(c,a)\,d\tau_c\,f(a)\,dx_a + \cdots \tag{6-23}$$

The first term of the series gives the wave function at the time t_b assuming the system to be free (or unperturbed, in case K_U is to be substituted for K_0) between the times t_a and t_b. Call this term ϕ. Thus

$$\phi(b) = \int K_0(b,a)f(a)\,dx_a \tag{6-24}$$

Using this definition, the series of Eq. (6-23) can be rewritten as

$$\psi(b) = \phi(b) - \frac{i}{\hbar}\int K_0(b,c)V(c)\phi(c)\,d\tau_c + \frac{i}{\hbar^2}\int\int K_0(b,c)V(c)K_0(c,d)V(d)\phi(d)\,d\tau_c\,d\tau_d + \cdots \tag{6-25}$$

In this form the series is called the *Born expansion* for ψ. If only the first two terms are included (thus only through first order in V), the result is the *first Born approximation*. It involves a single scattering by the potential V. This scattering occurs at the point c. Up to this point the system described by $\phi(c)$ is free; and after the scattering, the system moves from c to b, again free, and described by $K_0(b,c)$. An integral must be taken over all the possible points at which the scattering occurs. If three terms of the series are used (thus through terms of second order in V), the result is called the *second Born approximation*, etc.

---

† Note that our convention that $K(b,a)$ is zero for $t_b < t_a$ makes Eq. (6-22) invalid if $t_b < t_a$, but we shall not use it in this range of t values.

Problem 6-4 Using arguments similar to those leading to Eq. (6-19), show that the wave function $\psi(b)$ satisfies the integral equation

$$\psi(b) = \phi(b) - \frac{i}{\hbar} \int K_0(b,c) V(c) \psi(c) \, d\tau_c \tag{6-26}$$

This integral equation is equivalent to the Schrödinger equation

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} + \frac{\hbar^2}{2m}\nabla^2\psi + V\psi = 0 \tag{6-27}$$

Working in one dimension only, show how the Schrödinger equation may be deduced from the integral equation.

6-4 THE SCATTERING OF AN ELECTRON BY AN ATOM

Mathematical Treatment. We have developed the concepts and formulas of the perturbation treatment in a somewhat abstract framework. Now, to develop a physical understanding of the perturbation method, we shall discuss the specific problem of the scattering of a fast electron by an atom. We envision an experiment in which a beam of electrons bombards a target, such as a thin foil of metal, and then is collected by some suitable counter, as shown in Fig. 6-4.

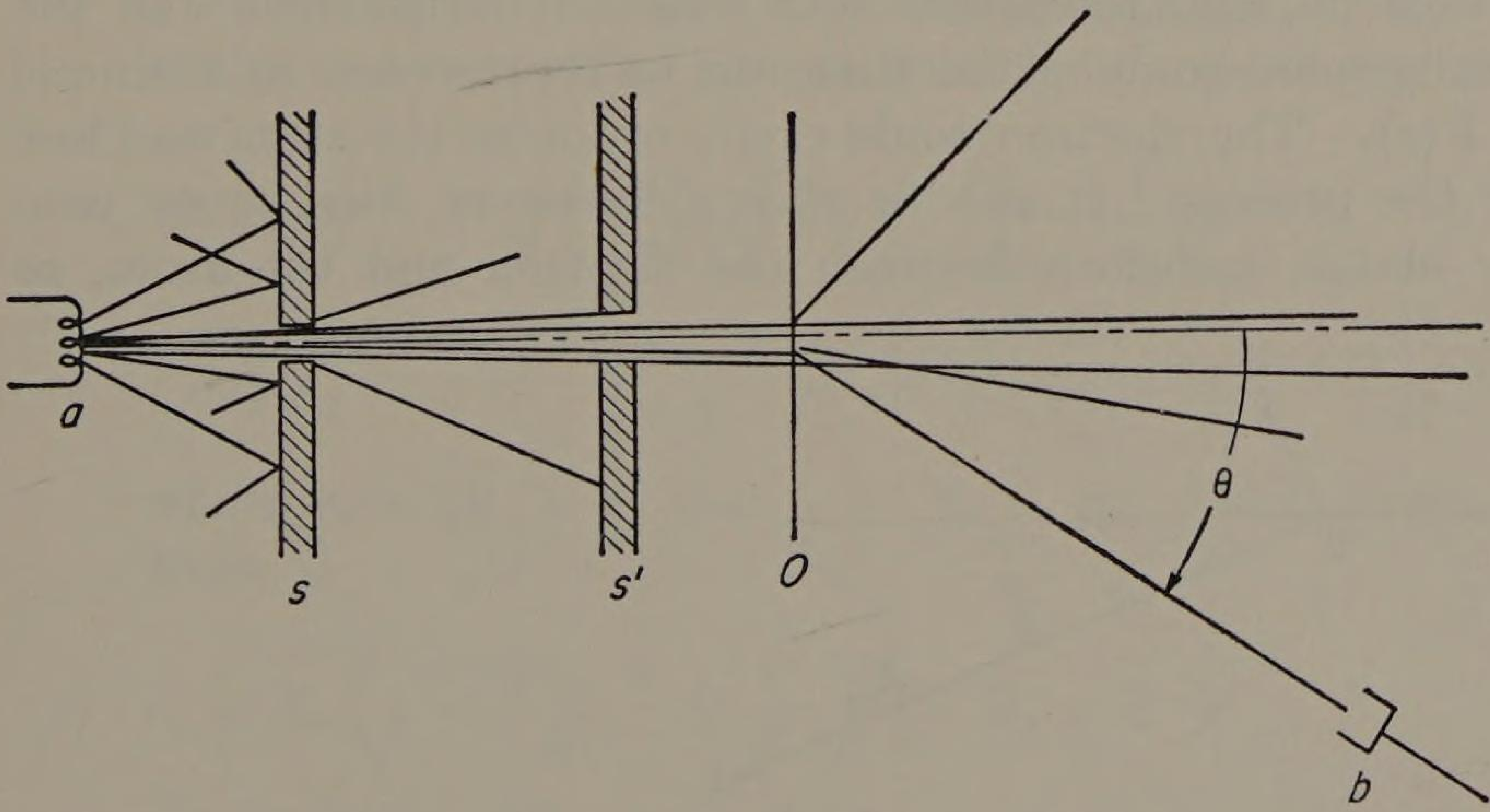


Fig. 6-4 Electrons boil off a hot filament at a, are screened into a beam by collimating holes in s and s', and then strike a thin-foil target at O. Most of the electrons pass straight on without being scattered (if their energy is great enough and the target is thin enough), but some are deflected by interactions with atoms in the target and scattered, for example, through an angle θ to b. As the counter at b is moved up and down, the relation between the relative number of scatterings and the scattering angle θ can be measured.

Suppose the energy of the scattering particles is determined by a time-of-flight method. That is, we release an electron from the source at one time, say, $t = 0$, and ask for the chance that it arrives at the counter after some delay T. We can then make direct use of our result for the amplitude $K(b,a)$ to go from one place to another in a definite time.

We shall simplify the problem by assuming that either the foil is so thin or the interaction is so weak that each electron can interact with, at most, one atom. Actually, this assumption is quite realistic for many scattering experiments. Furthermore, most multiple scatterings can be analyzed in terms of the simple scattering from one atom. Thus we shall discuss the interaction between a single electron and a single atom.

The center of the atom will be taken as the center of a coordinate system in which the electrons are released at the point a, as in Fig. 6-5, at the time $t = 0$. A counter placed at the point b tells us whether or not, at the time $t = T$, the electron arrives at the point b. We shall make the following approximations:

1. The interaction can be represented by a first-order Born approximation. That is, the electron is scattered only once by the atom.
2. The atom can be represented by a potential $V(\mathbf{r})$ fixed in space and constant in time.

Actually, the atom presents a very complicated system interacting with the electron, and the interaction between the electron and the atom is really more complicated than can be represented by a simple potential $V(\mathbf{r})$. The electron could excite or ionize the atom and lose energy in the process. It can be shown, however, that if we consider only elastic collisions between the electron and the atom, so

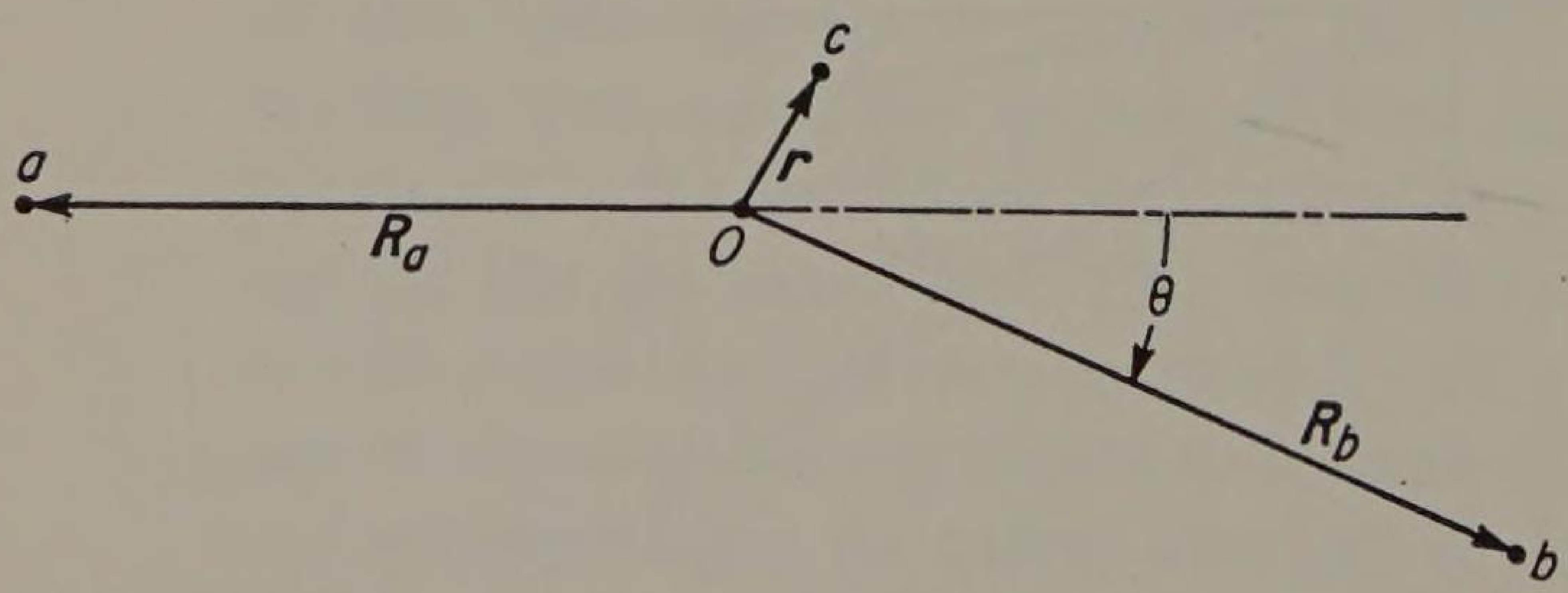


Fig. 6-5 The geometry of the scattering problem. The electron starts at a and moves as a free particle to c, where it is scattered by the atomic potential $V(r)$. After the scattering, it moves to the counter at b, which is located at the end of the radius vector $\mathbf{R}_b$ from the scattering center O. In this process, the electron has been scattered through the angle θ, measured from the direction of the nonscattered beam. This process corresponds to the first-order Born approximation. If the amplitude for two scatterings, say, at c and d, is included, then the result is the second-order approximation, etc.

that the atom is in the same energy state after the collision as it is before, then when the approximation (1) is valid, approximation (2) is valid too.

Let $\mathbf{R}_a$ and $\mathbf{R}_b$ be the vectors from the center of the atom to the points at which the electron is released and detected, respectively. In the calculations we shall take $\mathbf{R}_a$ and $\mathbf{R}_b$ to have lengths much larger than the radius of the atom. That is, we shall assume that the atomic potential $V(\mathbf{r})$ becomes negligibly small at distances much smaller than $|\mathbf{R}_a|$ and $|\mathbf{R}_b|$. Thus during most of its flight, the electron will be moving as a free particle, and only in the vicinity of the origin will it be exposed to the potential.

The first-order Born approximation contains two terms, only the second of which is of interest to us here. The first term is the kernel $K_0(b,a)$ for the motion of the electron from a to b as a free particle, and it has already been studied sufficiently. The term of interest is then

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\int K_0(b,c)V(c)K_0(c,a)\,d\tau_c$$

$$= -\frac{i}{\hbar}\int^{\mathbf{r}}\int_0^T \sqrt{\frac{m}{2\pi i\hbar(T-t)}}^{\,3} \left\{\exp\left[\frac{im|\mathbf{R}_a-\mathbf{r}|^2}{2\hbar(T-t)}\right]\right\} V(\mathbf{r})$$

$$\times \sqrt{\frac{m}{2\pi i\hbar t}}^{\,3}\left[\exp\left(\frac{im|\mathbf{R}_b-\mathbf{r}|^2}{2\hbar t}\right)\right] d^3\mathbf{r}\,dt \qquad (6\text{-}28)$$

Here we have used $\mathbf{r}$ as the vector from the origin to the point c, and $d^3\mathbf{r}$ represents the product of the differentials of all the components of the vector $\mathbf{r}$. The integral over t gives (cf. Appendix)

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\left(\frac{m}{2\pi i\hbar T}\right)^{5/2} T\int^{\mathbf{r}}\left(\frac{1}{r_a}+\frac{1}{r_b}\right)$$

$$\times\left\{\exp\left[\frac{im}{2\hbar T}(r_a+r_b)^2\right]\right\} V(\mathbf{r})\,d^3\mathbf{r} \qquad (6\text{-}29)$$

where $r_a = |\mathbf{R}_a - \mathbf{r}|$ and $r_b = |\mathbf{R}_b - \mathbf{r}|$. Using these definitions, we write

$$r_a = R_a\left(1 - \frac{2\mathbf{R}_a\cdot\mathbf{r}}{R_a^2} + \frac{r^2}{R_a^2}\right)^{1/2} \simeq R_a + \mathbf{i}_a\cdot\mathbf{r} \qquad (6\text{-}30)$$

$$r_b = R_b\left(1 - \frac{2\mathbf{R}_b\cdot\mathbf{r}}{R_b^2} + \frac{r^2}{R_b^2}\right)^{1/2} \simeq R_b - \mathbf{i}_b\cdot\mathbf{r} \qquad (6\text{-}31)$$

where $\mathbf{i}_a$ and $\mathbf{i}_b$ are unit vectors in the direction of $-\mathbf{R}_a$ and $\mathbf{R}_b$, respectively (that is, $\mathbf{i}_a = -\mathbf{R}_a/R_a$, where $R_a = |\mathbf{R}_a|$), and we have made use in the approximation of the fact that R_a is much larger than any value of $|\mathbf{r}|$ for which the potential is not negligible. It is necessary to keep the first-order terms in r only in the exponential factor, since this factor is quite sensitive to small relative changes in phase. Here we need

$$(r_a + r_b)^2 \simeq (R_a + R_b)^2 + 2(R_a + R_b)(\mathbf{i}_a \cdot \mathbf{r} - \mathbf{i}_b \cdot \mathbf{r}) \tag{6-32}$$

Using these approximations, the kernel can be written as

$$K^{(1)}(b,a) \simeq -\frac{i}{\hbar}\left(\frac{m}{2\pi i\hbar T}\right)^{5/2} T\left(\frac{1}{R_a} + \frac{1}{R_b}\right)\left\{\exp\left[\frac{im}{2\hbar T}(R_a + R_b)^2\right]\right\} \times \int^{\mathbf{r}} \left\{\exp\left[\frac{im}{\hbar T}(R_a + R_b)(\mathbf{i}_a \cdot \mathbf{r} - \mathbf{i}_b \cdot \mathbf{r})\right]\right\} V(\mathbf{r})\, d^3\mathbf{r} \tag{6-33}$$

Physical Interpretation. We can deduce some of the physical characteristics of the motion from a study of Eq. (6-33). In the time T the electron has traveled the total distance of $R_a + R_b$. Thus its velocity during this time is $u = (R_a + R_b)/T$ and its energy is $mu^2/2$, while its momentum has the magnitude mu. In writing these expressions we are making the assumption that the energy of the electron is not changed by the scattering process.

That these values for the velocities, energy, and momentum are consistent can be verified from an inspection of the exponential factor appearing in front of the integral of Eq. (6-33). The phase of this exponential term is $im(R_a + R_b)^2/2\hbar T$, and the derivative of this phase with respect to T gives the frequency as

$$\omega = \frac{m}{2\hbar}\frac{(R_a + R_b)^2}{T^2} \tag{6-34}$$

With u defined as above this means that the energy is $mu^2/2$ [cf. Eq. (3-15)].

Differentiating the phase with respect to R_b yields the wave number at the point b as

$$k = \frac{m}{\hbar}\frac{R_a + R_b}{T} \tag{6-35}$$

which means that the magnitude of the momentum is mu [cf. Eq. (3-12)].

Problem 6-5 The integral over t of Eq. (6-28) could be approximated by the method of stationary phase. By studying the application of such a method to this integral, show that most of the contribution to the integral comes from values of t near the region $t = R_a/u$, the time at which the electron would arrive at the center of the atom if it moved in a classical manner.

With the velocity of the electron defined as $u = (R_a + R_b)/T$ define the incoming vector of momentum $\mathbf{p}_a$ as

$$\mathbf{p}_a = mu\mathbf{i}_a \tag{6-36}$$

and the outgoing vector momentum $\mathbf{p}_b$ as

$$\mathbf{p}_b = mu\mathbf{i}_b \tag{6-37}$$

Then Eq. (6-33) can be written as

$$K^{(1)}(b,a) = -\frac{i}{\hbar}\left(\frac{m}{2\pi i\hbar}\right)^{5/2} \frac{u}{T^{1/2}R_aR_b}\left[\exp\left(\frac{im}{2\hbar}u^2T\right)\right] \times \int^{\mathbf{r}} \left\{\exp\left[\frac{i}{\hbar}(\mathbf{p}_a - \mathbf{p}_b)\cdot\mathbf{r}\right]\right\} V(\mathbf{r})\, d^3\mathbf{r} \tag{6-38}$$

Call the change in momentum, or the momentum transfer,

$$\mathbf{q} = \mathbf{p}_a - \mathbf{p}_b$$

and define the quantity $v(\mathbf{q})$ as

$$v(\mathbf{q}) = \int^{\mathbf{r}} e^{(i/\hbar)\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r})\, d^3\mathbf{r} \tag{6-39}$$

The probability that an electron arrives at the point b is given by the square of the absolute value of the kernel $K_V(b,a)$. Thus the probability will depend upon the first term in the series expansion of this kernel, namely, $K_0(b,a)$, which is likely to be so large as to completely overshadow the small perturbation term $K^{(1)}(b,a)$.

For this reason it is customary in most scattering experiments to collimate the incoming beam with suitable shields so that those electrons which are not scattered by the atoms in the target are confined to the region of a particular line (or direction), as shown in Fig. 6-6. Of course, there will be some diffraction by the collimating shields, such as that studied in Secs. 3-2 and 3-3, which means that some nonscattered electrons will appear outside this central beam. However, with suitable collimation, and for positions suitably far away from the collimated beam, the number of electrons diffracted by the collimator will be very small compared to the number scattered by the atoms in the target.

In such a region the probability of arrival for an electron is given, at least to first order, by the square of the absolute value of $K^{(1)}(b,a)$

alone. Using Eqs. (6-38) and (6-39), this probability is

$$\frac{P(b)}{\text{unit volume}} = \frac{1}{\hbar^2}\left(\frac{m}{2\pi\hbar}\right)^5 \frac{u^2}{TR_a{}^2R_b{}^2}\,|v(\mathbf{q})|^2 \tag{6-40}$$

In this last expression the factor $v(\mathbf{q})$ contains the characteristics of the atomic potential and the dependence of the kernel upon the relative directions of R_a and R_b. It is completely independent of the dimensions of the experimental equipment. The effects of such dimensions are represented by the remaining factors of Eq. (6-40). For example, the term $1/R_a{}^2$ can be easily seen to result from the idea that the chance for an electron to actually hit the atom varies inversely as $R_a{}^2$. The application of such an idea might be questioned in this experiment in view of the fact that we have supposed some collimating shields are present. However, this collimation has a negligible effect over *atomic* dimensions. From the point of view of a target atom the beam of oncoming electrons appears to consist of electrons spreading in all directions from a single source.

In a similar manner, after the scattering, the electrons spread out again in all directions from the scattering atom. Thus the chance per unit volume to find an electron in the counter varies inversely as $R_b{}^2$. Since the more interesting features of the experiment are contained in the function $v(\mathbf{q})$, we shall give special attention to this function in the next section.

The additional factors depend on the particular normalization of our kernel. We can interpret the formula more easily if we give it as a ratio. We compare the probability of finding a scattered particle at b to the probability of finding one at a point d behind the atom at the same total distance $R_a + R_b$ (and at the same time T, to keep the velocity the same) if no scattering occurred, as shown in Fig. 6-7. That is, we calculate $P(d)$ per unit volume, as if no atom were present.

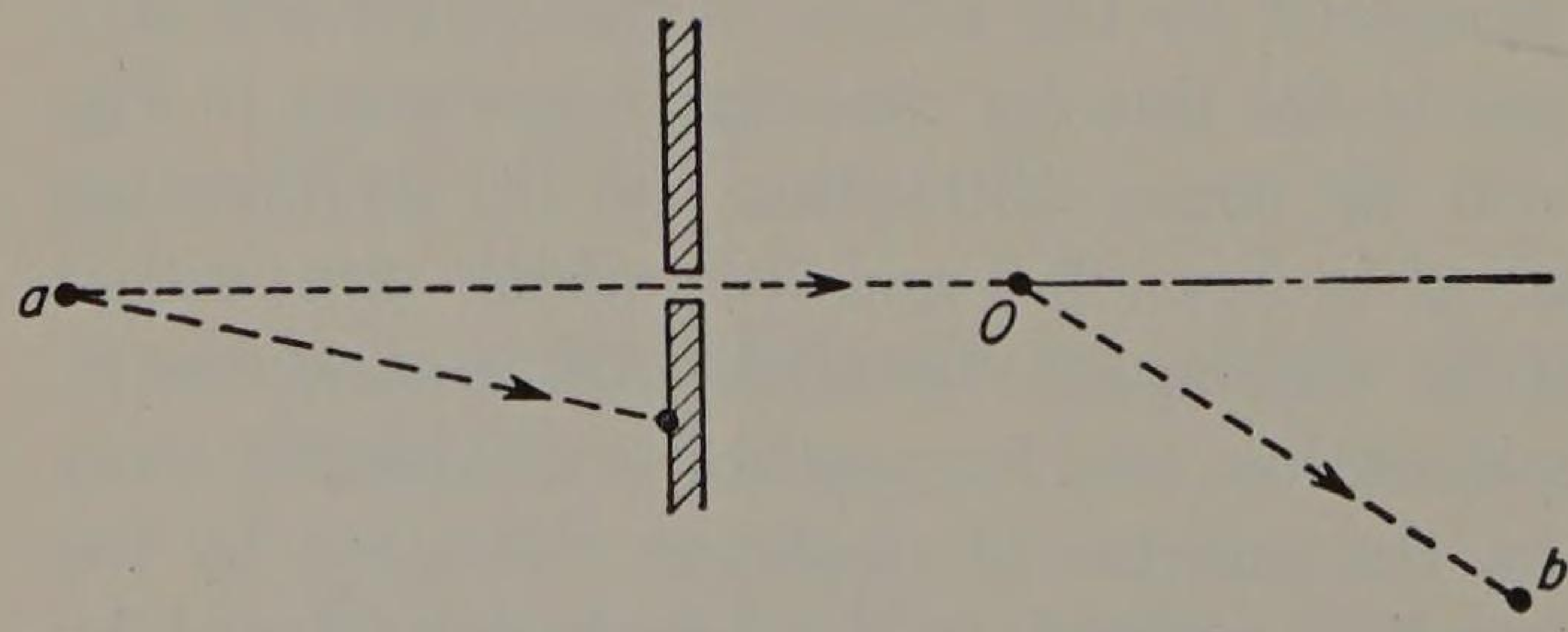


Fig. 6-6 Principle of collimation to eliminate the zero-order term at b. Only electrons which have been scattered at least once can get from a to b with any reasonable probability. Thus the zero-order term in the perturbation expansion of $K_V(b,a)$ will contribute a negligible amount and can be neglected. The first term of importance is $K^{(1)}(b,a)$.

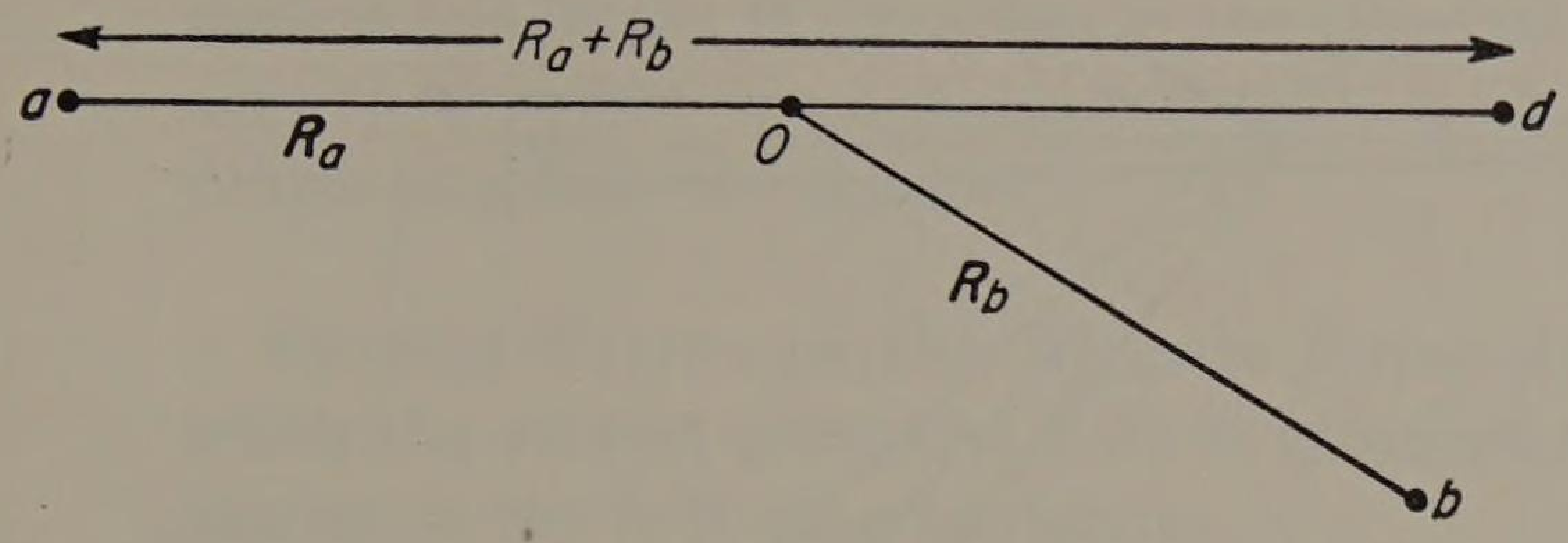


Fig. 6-7 If d and b are the same total distance from O, namely, R_b, then the difference (or ratio) of numbers of electrons arriving at the two points can depend only on the scattering phenomenon. If d is in the direct line of nonscattered electrons, the ratio of the number arriving at b to that arriving at d if no scattering source were present is the probability of scattering to b.

The result is $|K_0(d,a)|^2$ or

$$\frac{P(d)}{\text{unit volume}} = \left(\frac{m}{2\pi\hbar}\right)^3 \frac{u^2}{T(R_a+R_b)^2} \tag{6-41}$$

so that

$$\frac{P(b)}{P(d)} = \left(\frac{m}{2\pi\hbar^2}\right)^2 |v(\mathbf{q})|^2 \frac{(R_a+R_b)^2}{R_a{}^2R_b{}^2} \tag{6-42}$$

We shall interpret the last factor geometrically in the next section, where we shall also give more detailed attention to the function $v(\mathbf{q})$.

The Cross Section for Scattering. It is convenient to describe the characteristics of an atom in a scattering experiment by means of the concept of a *cross section.* The utility of such a concept stems from the convenience of thinking along the lines of classical physics. The cross section σ is defined as the effective target area (from a classical point of view) of the atom that must be hit by an electron in order that the electron be scattered into a unit solid angle. This solid angle is measured around a sphere whose center is at the atom. The cross section is thus a function of the scattering angle, i.e., the angle between $\mathbf{R}_a$ and $\mathbf{R}_b$. In terms of such a classical model we can determine the probability that an electron arrives at the point b.

If particles starting from the origin were to hit a small target of area $d\sigma$ at distance R_a, these particles would be removed from the region d, where they are spread out over an area $[(R_a+R_b)/R_a]^2\,d\sigma$. Instead they are sent out in a solid angle $d\Omega$ toward b and are therefore spread out over an area $R_b{}^2\,d\Omega$ there, as shown in Fig. 6-8. Hence the ratio of the probability of finding them at b to that of finding them

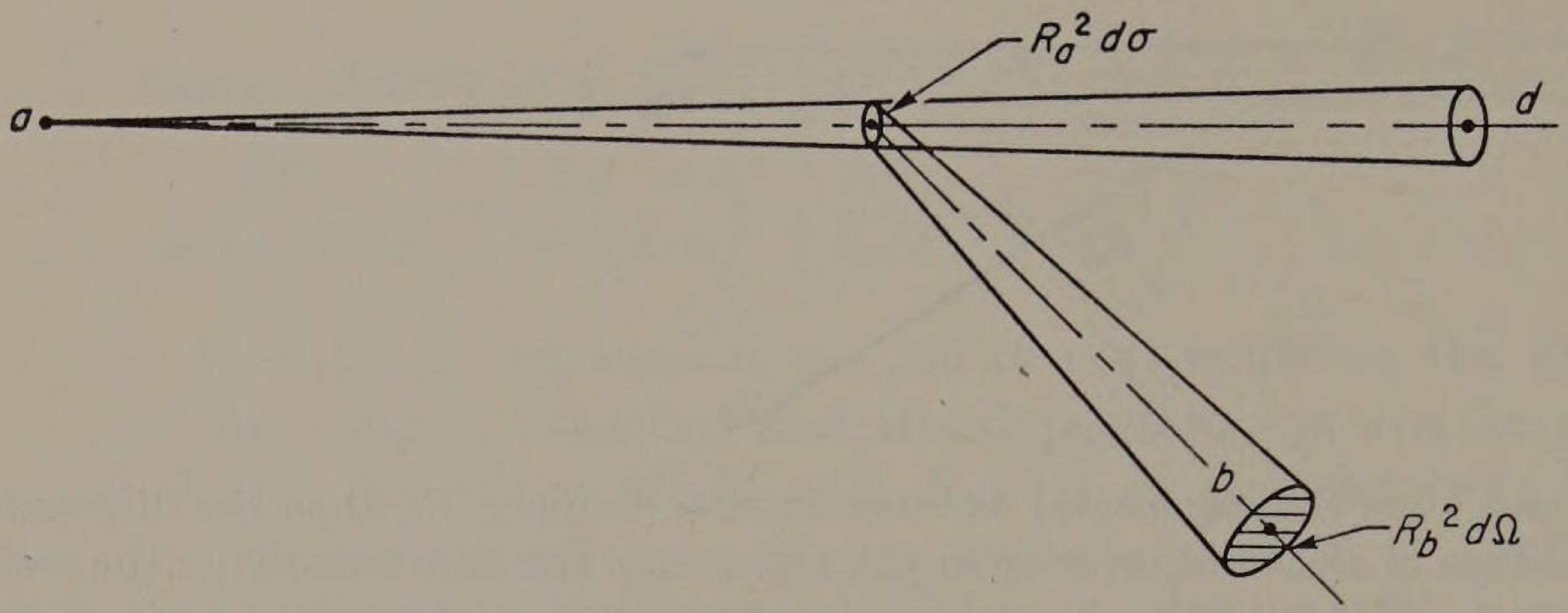


Fig. 6-8 Particle striking an area $d\sigma$ of the target are deflected through an angle θ into an area measured by the solid angle $d\Omega$. If no target had been present, the particles would have proceeded to the point d. Instead, they proceed to the point b, spreading out into the area $R_b^2\,d\Omega$. The probability of finding a particle at d is inversely proportional to the area over which the beam would have spread in arriving at d. Similarly, the probability of finding the particle at b is inversely proportional to the area $R_b^2\,d\Omega$ over which the beam of scattered particles spreads in traveling from the target to b. If we take the ratio of these areas, we have the inverse ratio of the associated probabilities. From this point of view we say that all of the particles which hit the target area $d\sigma$ are scattered, and through a particular angle θ. Of course, actually only a few particles which hit the target are scattered at all and only a fraction of these through the angle θ. Thus, the area element $d\sigma$ which we have used in this calculation is the *effective* cross-sectional area for scattering through the angle θ measured in terms of the element of solid angle $d\Omega$ into which the particles are scattered.

at d if there were no target is the inverse ratio of these areas,

$$\frac{P(b)}{P(d)} = \frac{(R_a + R_b)^2\,d\sigma/R_a^2}{R_b^2\,d\Omega} \tag{6-43}$$

On comparing Eqs. (6-42) and (6-43), we see that the cross section per unit solid angle is

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 |v(\mathbf{q})|^2 \tag{6-44}$$

The main advantage of an expression in terms of cross section instead of using Eq. (6-40) directly is this: Equation (6-44) does not depend on particular experimental conditions, so the cross sections obtained in one or another experiment can be directly compared, whereas probabilities per unit volume cannot be.

It must be emphasized that this idea of an effective target is purely classical and is convenient in recording scattering probabilities. There is no direct relation between it and the size of the atom, nor is the scattering mechanism to be thought of as localized exactly over such an area. For example, the shadow one would expect to find classically behind the target is not there in the classical sense (with sharp boundaries); for since we are dealing with a wave phenomenon, there is diffraction into the shadow.

Special Forms of the Atomic Potential. The results obtained when the atomic potential $V(\mathbf{r})$ is assumed to have various forms are shown in the following problems.

Problem 6-6 Suppose the potential is that of a central force. Thus $V(\mathbf{r}) = V(r)$. Show that $v(\mathbf{q})$ can be written as

$$v(\mathbf{q}) = v(q) = \frac{4\pi\hbar}{q} \int_0^\infty r \left(\sin \frac{qr}{\hbar}\right) V(r)\, dr \tag{6-45}$$

Suppose $V(r)$ is the Coulomb potential Ze^2/r. In this case the integral for $v(q)$ is oscillatory at the upper limit. But convergence of the integral can be artificially forced by introducing the factor $e^{-\epsilon r}$ and then taking the limit of the result as $\epsilon \to 0$. Following through this calculation, show that the cross section corresponds to the Rutherford cross section

$$\sigma_{Ruth} = \frac{4Ze^4m^2}{q^4} = \frac{Ze^4}{16(mu^2/2)[\sin(\theta/2)]^4} \tag{6-46}$$

where e = charge on an electron

$q = 2p \sin(\theta/2) = 2mu \sin(\theta/2)$ (6-47)

θ = angle between the vectors $\mathbf{i}_a$ and $\mathbf{i}_b$

The result of Prob. 6-6 is, accidentally, exact. That is, the first-order Born approximation gives the exact value for the probability of scattering in a Coulomb potential. This does not mean the higher-order terms are zero; it means, rather, that they contribute only to the phase of the scattering amplitude. Since the probability is the absolute square of the amplitude, it is independent of the phase. Thus a first-order Born approximation, which gives the correct value for the probability, is not exact for the amplitude. This case of a Coulomb scattering is amusing, for there is also another accident. A completely classical treatment of such a scattering problem, i.e., treating the electrons as charged point masses, gives the same result.

Problem 6-7 Suppose the potential $V(\mathbf{r})$ is the result of a charge distribution $\rho(\mathbf{r})$ so that

$$\nabla^2 V(\mathbf{r}) = 4\pi e \rho(\mathbf{r}) \tag{6-48}$$

By assuming that $\rho(\mathbf{r})$ goes to 0 as $|\mathbf{r}| \to \infty$, multiplying Eq. (6-48) by $\exp(i\mathbf{q}\cdot\mathbf{r}/\hbar)$, and integrating twice over $\mathbf{r}$, show that $v(\mathbf{q})$ can be expressed in terms of ρ as

$$v(\mathbf{q}) = \frac{4\pi\hbar^2 e^2}{q^2} \int^{\mathbf{r}} e^{(i/\hbar)\mathbf{q}\cdot\mathbf{r}} \rho(\mathbf{r})\, d^3\mathbf{r} \tag{6-49}$$

An atom can be represented in terms of its charge density. At the nucleus the charge density is singular, so that it can be represented as a δ function of r of strength Z, where Z is the charge on the nucleus. Then if ρ_e is the density of atomic electrons, $v(\mathbf{q})$ is

$$v(\mathbf{q}) = \frac{4\pi\hbar^2 e^2}{q^2} \left[Z - \int^{\mathbf{r}} \rho_e(\mathbf{r}) e^{(i/\hbar)\mathbf{q}\cdot\mathbf{r}}\, d^3\mathbf{r} \right] \tag{6-50}$$

The quantity in the brackets is called the *form factor* for electron scattering. (Incidentally, a similar form factor appears in X-ray scattering. The theory of X-ray scattering shows that only the atomic electrons, and not the nucleus, contribute to the scattering. Thus the form factor for X-ray scattering is the same but with the Z omitted.)

In an atom the potential follows the Coulomb law only for very small radii. As the radius is increased the atomic electrons gradually shield, or cancel out, the nuclear charge until, for sufficiently large values of r, the potential is zero. The shielding effect of atomic electrons can be accounted for in a very rough approximate manner with the formula

$$V(r) = \frac{Ze^2}{r} e^{-(r/a)} \tag{6-51}$$

In this expression a is called the radius of the atom. It is not the same as the outer radius of the atom as used by chemists, but instead is given by $a_0/Z^{1/3}$, where $a_0 = \hbar^2/me^2 = 0.528$ Å.

Problem 6-8 Show that in such a potential

$$v(q) = \frac{4\pi Z e^2 \hbar^2}{q^2 + (\hbar/a)^2} \tag{6-52}$$

and hence

$$\sigma = Z^2 e^4 \left\{ \frac{mu^2}{2} \left[4 \left(\sin\frac{\theta}{2} \right)^2 + \frac{\hbar^2}{(pa)^2} \right] \right\}^{-2} \tag{6-53}$$

The total cross section σ_T is defined as the integral of σ over the unit sphere; thus

$$\sigma_T = \int_0^{4\pi} \sigma\, d\Omega \tag{6-54}$$

In the present example show that σ_T is given by

$$\sigma_T = \pi a^2 \frac{Z^2e^4/(2u\hbar)^2}{1 + \hbar^2/(2pa)^2} \tag{6-55}$$

Problem 6-9 Suppose we introduce the fact that the atomic nucleus has a finite radius given by

$$r = 1.2 \times 10^{-13} \times (\text{mass number})^{1/3}\ \text{cm} \tag{6-56}$$

and assume that the nuclear charge is distributed approximately uniformly in a sphere of this radius. What is the effect of this assumption on the cross section for the scattering of electrons by atoms at large values of the momentum transfer q?

Show how the nuclear radius can be determined along with some of the details of the nuclear charge distribution by making use of this effect. How large must the momentum p of the incoming electrons be in order to produce an appreciable effect? Would one observe more carefully the large or small scattering angles? Why?

Note: In this type of experiment the required electron momentum is so high that actually the relativistic formula $E = \sqrt{m^2c^4 + c^2p^2} - mc^2$ must be used to find the energy. So, strictly, we should not be allowed to use nonrelativistic formulas to describe the interaction. However, the relations between momentum and wavelength and between energy and frequency are not changed in the relativistic region. Since it is the wavelength which determines the resolving power of this "electron microscope," the momentum calculated by nonrelativistic formulas is still correct.

Problem 6-10 Consider a diatomic molecule containing two atoms, A and B, arranged with their centers at the points given by the vectors **a** and **b**. Using the Born approximation, show that the amplitude for an electron to be scattered from such a molecule is

$$K^{(1)} = e^{(i/\hbar)\mathbf{q}\cdot\mathbf{a}}f_A(\mathbf{q}) + e^{(i/\hbar)\mathbf{q}\cdot\mathbf{b}}f_B(\mathbf{q}) \tag{6-57}$$

where f_A and f_B are the amplitudes for scattering by the two atoms individually when each atom is located at the center of a coordinate system. The atomic binding does not change the charge distributions around the nuclei very much (except for very light nuclei such as hydrogen) because the binding forces affect only a few of the outermost electrons.

Using Eq. (6-57), show that the probability of scattering at a particular value of q is proportional to $f_A^2 + f_B^2 + 2f_Af_B \cos(\mathbf{q} \cdot \mathbf{d})$, where $\mathbf{d}$ is $\mathbf{a} - \mathbf{b}$.

The values of f computed by the Born approximation are real and the result is valid for electron energies usually used in diffraction experiments on molecules (the order of 1 kev). However, if the molecule includes the very heaviest atoms, such as uranium, the atomic potential is too large for the results to be adequately described by the Born approximation, and small corrections are necessary.

Problem 6-11 Suppose the molecules are oriented in a random fashion. Show that the electron scattering averaged over a group of such molecules is proportional to $f_A^2 + f_B^2 + 2f_Af_B (\sin \mathbf{q} \cdot \mathbf{d})/\mathbf{q} \cdot \mathbf{d}$. How can this result be generalized to the case of polyatomic molecules?

These results form the basis of electron diffraction techniques which make possible the determination of the form of molecules.

Problem 6-12 Assume that $V(\mathbf{r})$ is independent of time and show that the time integral of the second-order scattering term $K^{(2)}(b,a)$ gives

$$K^{(2)}(b,a) = \left(\frac{m}{2\pi\hbar^2}\right)^2 \left(\frac{m}{2\pi i\hbar T}\right)^{3/2} \int^{\mathbf{r}_c} \int^{\mathbf{r}_d} \frac{r_{cd} + r_{ac} + r_{bd}}{r_{cd}r_{ac}r_{bd}} \times \left[\exp\left(\frac{im}{2\hbar T}\right)(r_{cd} + r_{ac} + r_{bd})^2\right] V(\mathbf{r}_c)V(\mathbf{r}_d)\, d^3\mathbf{r}_c\, d^3\mathbf{r}_d \tag{6-58}$$

where the points a, b, c, and d are arranged as shown in Fig. 6-9. The term r_{cd} stands for the distance between the point c and d, etc.

Assume that $V(\mathbf{r})$ becomes negligibly small at distances which are short compared to R_a or R_b. Show that the cross section is given by $\sigma = |f|^2$, where the scattering amplitude f, including the first-order term, is

$$f = \frac{m}{2\pi\hbar^2} \int^{\mathbf{r}} e^{-(i/\hbar)\mathbf{p}_b\cdot\mathbf{r}} V(\mathbf{r}) e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{r}}\, d^3\mathbf{r} + \left(\frac{m}{2\pi\hbar^2}\right)^2 \int^{\mathbf{r}_c} \int^{\mathbf{r}_d} e^{-(i/\hbar)\mathbf{p}_b\cdot\mathbf{r}_d} V(\mathbf{r}_d) \left(\frac{1}{r_{cd}}\right) e^{(i/\hbar)pr_{cd}} \cdot V(\mathbf{r}_c) e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{r}_c}\, d^3\mathbf{r}_c\, d^3\mathbf{r}_d + \text{higher-order terms} \tag{6-59}$$

where $\mathbf{p}_b$ is the momentum of the electron traveling in the direction of $\mathbf{R}_b$ and $\mathbf{p}_a$ is the momentum of an electron traveling in the direction

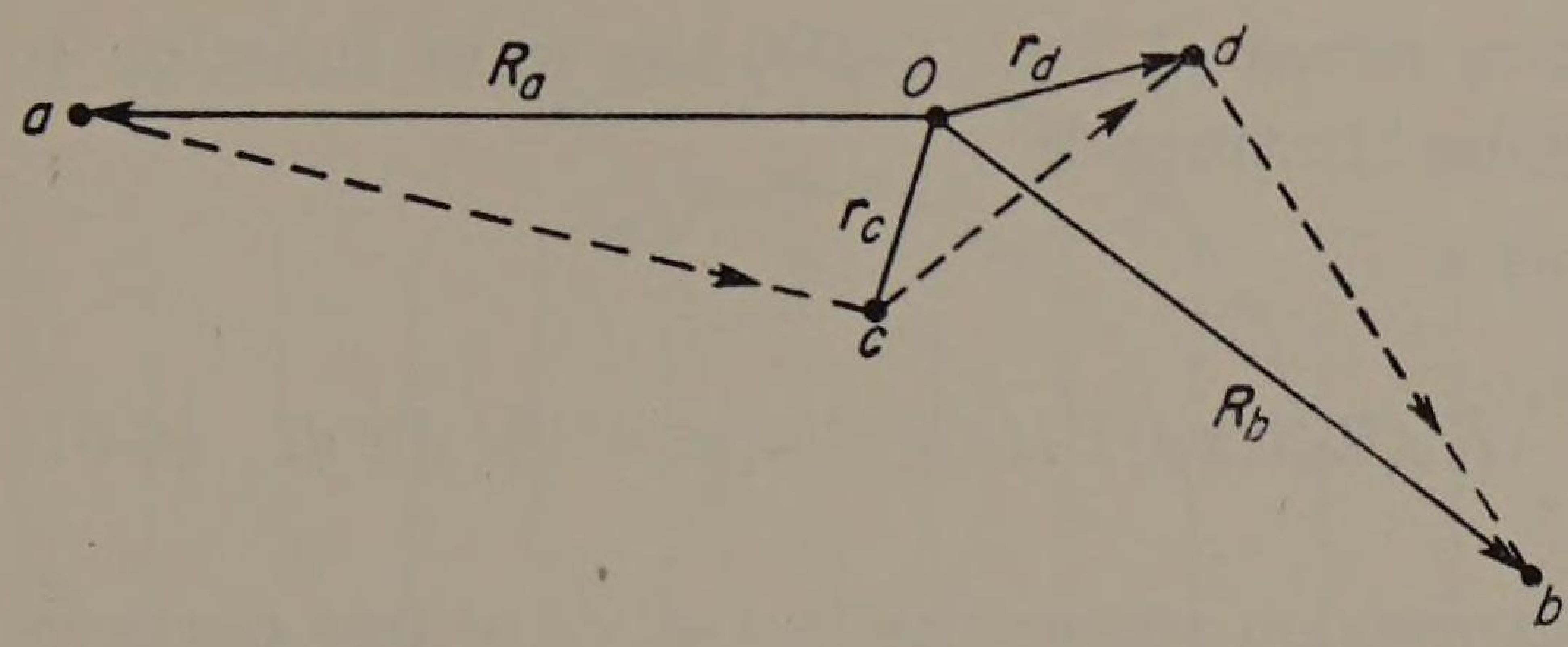


Fig. 6-9 To increase the accuracy of scattering calculations, we can take account of second-order terms in the perturbation expansion. Here, as in Fig. 6-2 (3), we picture the electron as being scattered at two separate points in the atomic potential. Thus the electron starts at a; proceeds as a free particle to c, where it is scattered; then moves as a free particle to d, where it is scattered again; and finally moves as a free particle to b, where it is collected by the counter. The points c and d can lie at any position in space. The atomic potential at these positions depends upon the radius vectors $\mathbf{r}_c$ and $\mathbf{r}_d$, measured from the center of the atom O.

$-\mathbf{R}_a$. The magnitude of the momentum is p, and it is approximately unchanged by an elastic scattering of the electron from the (relatively massive) atom.

One might expect that in a situation in which the Born approximation is not adequate it would be worthwhile to compute the second-order term as a correction. But in practice it seems that in this application Eq. (6-59) is a kind of asymptotic series. If the second term makes an appreciable correction (say 10 per cent or more) the higher terms are not much smaller and the true correction cannot be gotten easily by this method. Of course, if it is a problem in which the errors of the Born approximation are small (say less than 1 per cent), the second term will be adequate to find the corrections.

The Wave Function Treatment of Scattering. In the scattering experiment which we have described we have assumed that the initial state of the incoming electron was that of a free particle with momentum $\mathbf{p}_a$. We have assumed that the value of the momentum is determined by a time-of-flight technique (i.e., the total time required to travel the distance $R_a + R_b$ is T).

It is not necessary to use such a technique. Any device which enables us to determine the momentum is equally satisfactory. So suppose we generalize our picture of scattering phenomena with the help of the wave function method.

Suppose the incoming electrons are known to have momentum p_a and energy $E_a = p_a^2/2m$. Thus the wave function for the incoming electrons is

$$\phi_a = e^{(i/\hbar)\mathbf{p}_a \cdot \mathbf{r}} e^{-(i/\hbar)E_a t} \tag{6-60}$$

Then, using the first two terms of Eq. (6-25), the wave function to first order for the outgoing electrons is

$$\psi(\mathbf{R}_b,t_b) = e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{R}_b}e^{-(i/\hbar)E_at_b} - \frac{i}{\hbar}\int_0^{t_b}\int^{\mathbf{r}} K_0(\mathbf{R}_b,t_b;\mathbf{r},t)\,V(\mathbf{r},t)e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{r}}e^{-(i/\hbar)E_at}\,d^3\mathbf{r}\,dt \qquad (6\text{-}61)$$

The first term in this expression represents the wave of free particles which have passed through the potential region without being scattered. The second term is the amplitude of the scattered electrons. If ψ_s represents this term, then ψ_s is the scattered wave.

Problem 6-13 Assume $V(\mathbf{r},t)$ is actually independent of t. Substituting for the free-particle kernel K_0 in Eq. (6-61), integrate the result over t to show that

$$\psi(\mathbf{R}_b,t_b) = e^{(i/\hbar)E_bt_b}\left[e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{R}_b} + \frac{m}{2\pi\hbar^2}\int^{\mathbf{r}_c}\frac{1}{r_{bc}}\,e^{(i/\hbar)pr_{bc}}V(\mathbf{r}_c)e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{r}_c}\,d^3\mathbf{r}_c\right] \qquad (6\text{-}62)$$

where r_{bc} is the distance from the final point b to the variable point of integration c and p is the magnitude of the momentum of the electron.

Once more, suppose that the potential drops to 0 for distances which are short compared to either R_a or R_b. Show that Eq. (6-62) can be written as

$$\psi(\mathbf{R}_b,t_b) = e^{(i/\hbar)E_bt_b}e^{(i/\hbar)\mathbf{p}_a\cdot\mathbf{R}_b} + f\,\frac{e^{(i/\hbar)pR_b}}{R_b} \qquad (6\text{-}63)$$

where the scattering amplitude f is defined in terms of $v(\mathbf{q})$ [see Eq. (6-39)] as

$$f = \frac{m}{2\pi\hbar^2}\,v(\mathbf{q}) \qquad (6\text{-}64)$$

The last term of Eq. (6-63), $(f/R_b)\exp(ipR_b/\hbar)$, can be thought of as the spatial part of the scattered wave function. It has the form of a spherical wave radiating outward from the center of the scattering atom. The amplitude of this spherical wave at some particular scattering angle depends upon that angle through the function f, which, by Eq. (6-64), varies with the momentum transfer q. Thus the complete wave functions for the electrons after scattering can be thought of as the sum of two terms. The first term is the plane wave

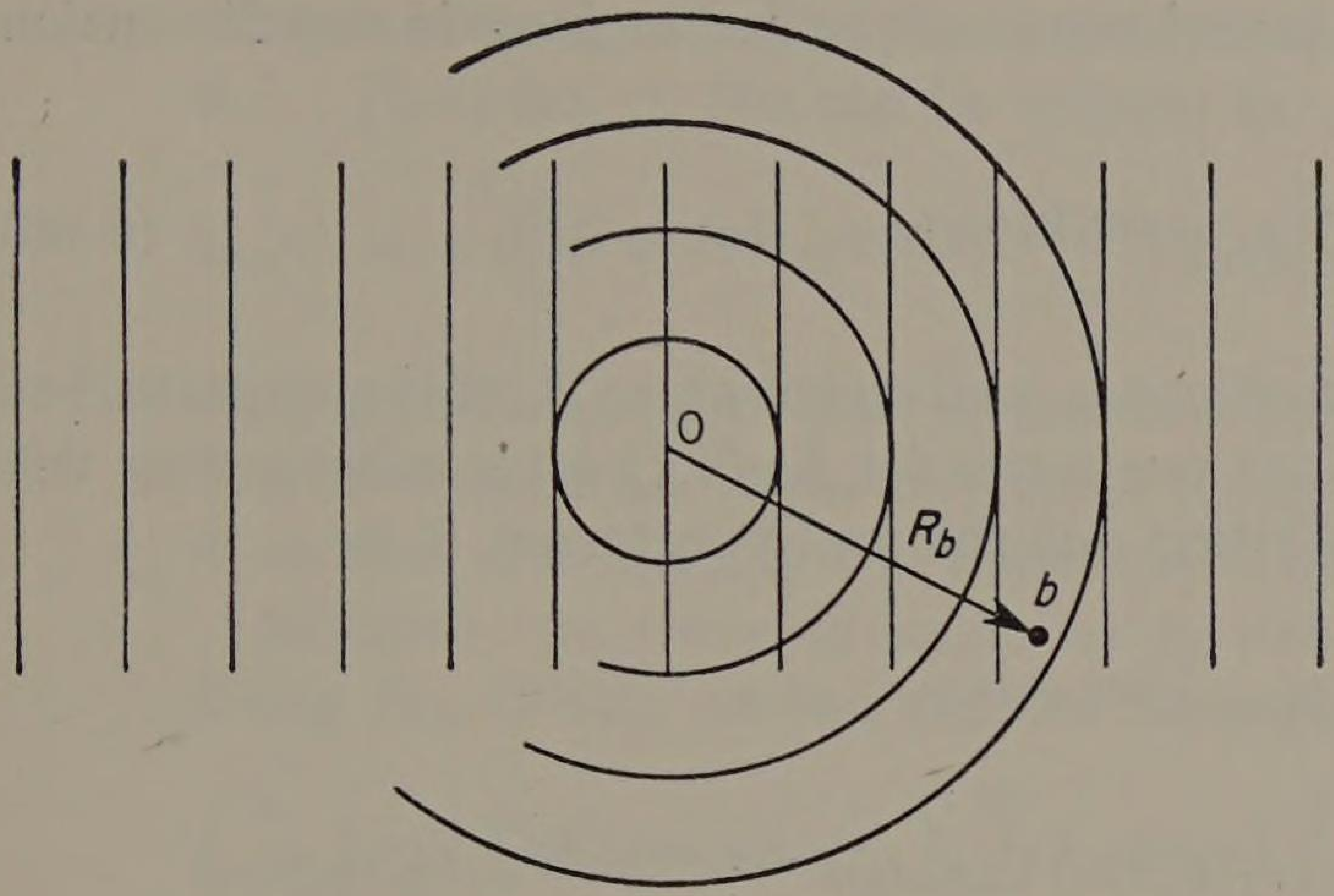


Fig. 6-10 A beam of electrons, represented by its equivalent wave, moves toward the atomic nucleus at O. The strongest fraction of the beam moves on undisturbed as a plane wave with momentum p_a. A small fraction of the beam is scattered from the atomic nucleus and moves away from O as a spherical wave pattern. The resulting strength of the wave of electrons (the number of electrons) at some point b located at R_b, measured from the atomic nucleus O, is then made up of two parts. The first is the nonscattered beam given by the plane wave $\exp (i\mathbf{p}_a \cdot \mathbf{R}_b/\hbar)$. To this is added the scattered wave with the spherical form given by $(1/R_b) \exp (ipR_b/\hbar)$ and multiplied by the function f, which determines the angular dependence. The combination of these two waves gives the spatial part of the scattered wave function.

of nonscattered electrons, $\exp (i\mathbf{p}_a \cdot \mathbf{R}_b/\hbar)$, and the second term is the spherical wave of scattered electrons, as indicated in Fig. 6-10. Use this point of view to derive the formula for the cross section σ.

Problem 6-14 Use the wave function approach to discuss the scattering of an electron from a sinusoidally oscillating field whose potential is given by

$$V(r,t) = U(r) \cos \omega t \tag{6-65}$$

Show that in the first-order Born approximation the energy of the outgoing wave is changed by either $+\omega$ or $-\omega$. What happens in the higher-order terms?

6-5 TIME - DEPENDENT PERTURBATIONS AND TRANSITION AMPLITUDES

The Transition Amplitude. An especially useful form of the perturbation theory occurs if the unperturbed problem corresponds to a potential U independent of time, for then we have seen in Eq. (4-59) that the unperturbed kernel can be expanded as (now in one dimension for convenience)

$$K_U(2,1) = \sum_n \phi_n(x_2)\phi_n^*(x_1)e^{-(iE_n/\hbar)(t_2-t_1)} \qquad \text{for } t_2 > t_1 \tag{6-66}$$

in terms of the eigenfunctions ϕ_n and eigenvalues E_n of the unperturbed problem. Let us look at our series for $K_V(2,1)$ after substituting this expression for K_U. Writing out the first two terms, it is

$$\begin{aligned} K_V(2,1) = {} & \sum_n \phi_n(x_2)\phi_n^*(x_1)e^{-(iE_n/\hbar)(t_2-t_1)} \\ & - \frac{i}{\hbar}\sum_n\sum_m \int \phi_m(x_2)\phi_m^*(x_3)V(x_3,t_3)e^{-(iE_n/\hbar)(t_2-t_3)}\phi_n(x_3)\phi_n^*(x_1) \\ & \qquad \times e^{-(iE_n/\hbar)(t_3-t_1)}\,dx_3\,dt_3 + \cdots \end{aligned} \tag{6-67}$$

It is clear that in each term the x_1 will appear in some wave function, like $\phi_m^*(x_1)$, and the x_2 likewise, so we can always write K_V in the form

$$K_V(2,1) = \sum_n\sum_m \lambda_{mn}(t_2,t_1)\phi_m(x_2)\phi_n^*(x_1) \tag{6-68}$$

where the λ's are coefficients depending on t_2, t_1. We shall call these coefficients *transition amplitudes*. To zero order in V, this must reduce to K_U, so to this order $\lambda_{mn} = \delta_{mn}e^{-(iE_n/\hbar)(t_2-t_1)}$. If we expand λ in a series in increasing orders of V, we have

$$\lambda_{mn} = \delta_{mn}e^{-(iE_n/\hbar)(t_2-t_1)} + \lambda_{mn}{}^{(1)} + \lambda_{mn}{}^{(2)} + \cdots \tag{6-69}$$

and comparison to Eq. (6-67) shows

$$\lambda_{mn}{}^{(1)} = -\frac{i}{\hbar}\int_{-\infty}^{\infty}\int_{t_1}^{t_2} \phi_m^*(x_3)V(x_3,t_3)\phi_n(x_3)\,dx_3\,e^{(i/\hbar)[E_m(t_3-t_2)-E_n(t_3-t_1)]}\,dt_3 \tag{6-70}$$

Problem 6-15 Recall that in Prob. 5-4 we defined a particular integral as the transition amplitude to go from state $\psi(x)$ to state $\chi(x)$. Show that the function λ_{mn} satisfies this definition when the initial state is the eigenfunction $\phi_n(x)$ and the final state is the eigenfunction $\phi_m(x)$.

Define, for brevity,

$$V_{mn}(t_3) = \int_{-\infty}^{\infty} \phi_m^*(x_3)V(x_3,t_3)\phi_n(x_3)\,dx_3 \tag{6-71}$$

(This is sometimes called the *matrix element* of V between states n and m.) Then Eq. (6-70) can be written as

$$\lambda_{mn}{}^{(1)} = -\frac{i}{\hbar} e^{-(i/\hbar)E_m t_2} e^{+(i/\hbar)E_n t_1} \int_{t_1}^{t_2} V_{mn}(t_3) e^{(i/\hbar)(E_m - E_n)t_3}\, dt_3 \tag{6-72}$$

This is an important result of the time-dependent perturbation theory. The coefficient λ_{mn} is the amplitude for the system to be found in state m at time t_2 if initially the system is in state n.

Suppose the wave function at t_1 was $\phi_n(x_1)$. What is it at t_2? Using Eq. (3-42), we can express the wave function at t_2 as

$$\int_{-\infty}^{\infty} K_V(2,1)\phi_n(x_1)\, dx_1 = \sum_k \sum_l \lambda_{kl}\phi_k(x_2) \int_{-\infty}^{\infty} \phi_l^*(x_1)\phi_n(x_1)\, dx_1$$
$$= \sum_k \lambda_{kn}\phi_k(x_2) \tag{6-73}$$

That is, the wave function at t_2 is in the form $\sum_m C_m\phi_m(x_2)$.

This expansion in terms of eigenfunctions was first introduced in Eq. (4-48). Now we can assign a deeper meaning to the constants C_m. We can interpret C_m as the amplitude that the system is in state ϕ_m. In this particular case, $C_m = \lambda_{mn}$ is the amplitude to be in the state ϕ_m at time t_2 if at t_1 the state is ϕ_n.

With no potential acting, a system once in state n is always in state n with an amplitude varying in time. So, to zero order,

$$\lambda_{mn} = \delta_{mn} e^{-(iE_n/\hbar)(t_2 - t_1)}$$

We can interpret the first-order term by the rule (Fig. 6-11): *The amplitude to be scattered from state n to m in a time dt is $-(i/\hbar)V_{mn}(t)\, dt$.*

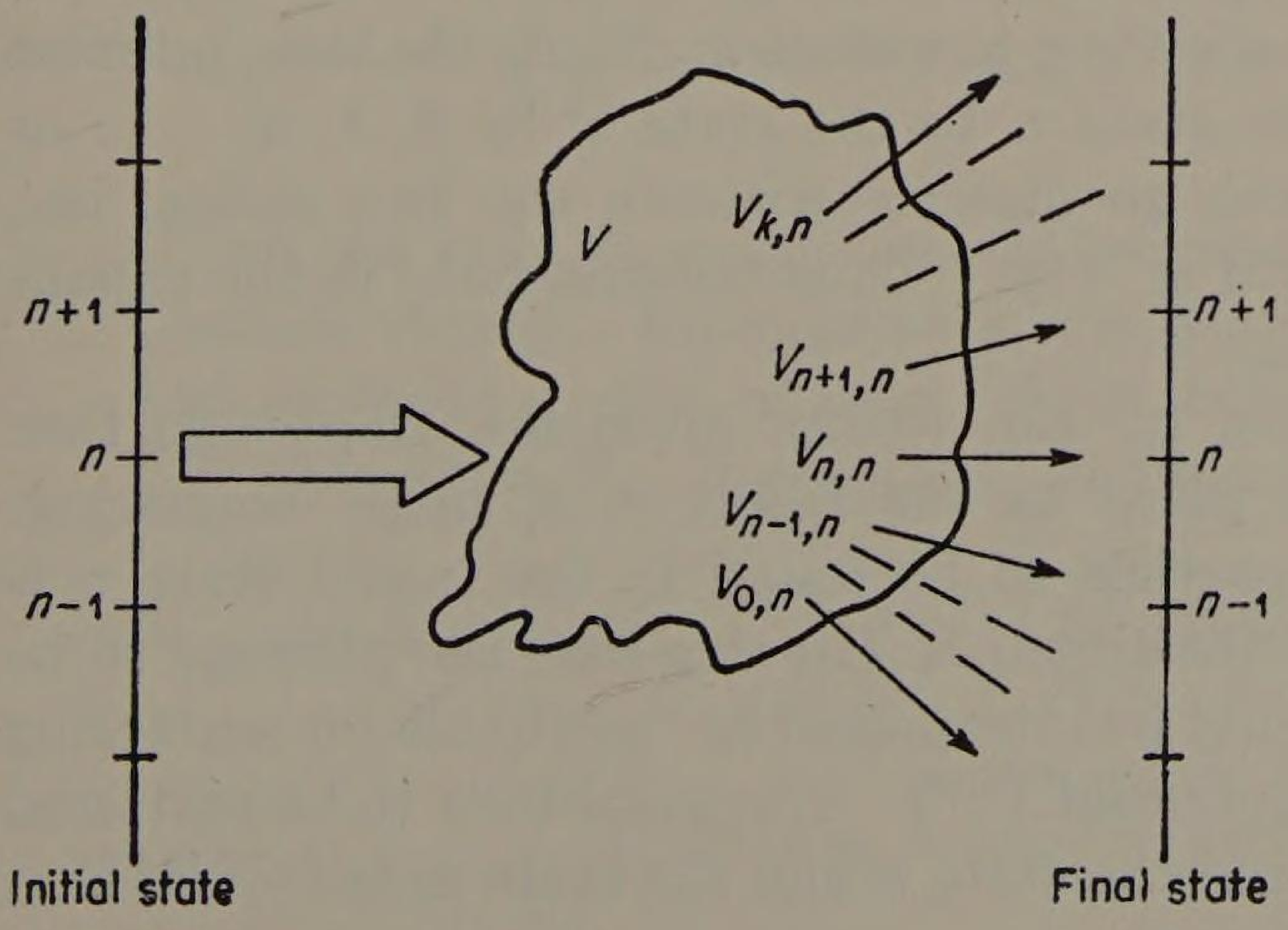


Fig. 6-11 A system initially at the nth energy level is subjected to a potential V which "scatters" the system into all of the states available to it. The amplitude for a scattering into the kth state, for example, is proportional to V_{kn}. In particular, the amplitude to be scattered from the state n to the state m in the time interval dt is $-(i/\hbar)V_{mn}(t)\, dt$.

Problem 6-16 Interpret Eq. (6-71) as a sum over alternatives; i.e., identify the alternatives.

Problem 6-17 Interpret Eq. (6-72) by explaining the meaning of each term. Then explain and verify the equation for the second-order coefficient

$$\lambda_{mn}^{(2)} = -\frac{1}{\hbar^2}\int_{t_1}^{t_2}\left[\int_{t_1}^{t_4}\sum_k e^{-(i/\hbar)E_m(t_2-t_4)}V_{mk}(t_4)\right.$$
$$\left.\times\, e^{-(i/\hbar)E_k(t_4-t_3)}V_{kn}(t_3)e^{-(i/\hbar)E_n(t_3-t_1)}\,dt_3\right]dt_4 \qquad (6\text{-}74)$$

Problem 6-18 Derive and interpret the integral equation

$$\lambda_{mn}(t_2,t_1) = \delta_{mn}e^{-(i/\hbar)E_m(t_2-t_1)}$$
$$-\frac{i}{\hbar}\int_{t_1}^{t_2} e^{-(i/\hbar)E_m(t_2-t_3)}\sum_k V_{mk}(t_3)\lambda_{kn}(t_3,t_1)\,dt_3 \qquad (6\text{-}75)$$

Problem 6-19 Consider $\lambda_{mn}(t_2)$ as a function of the final time t_2. Show, using either Eq. (6-75) or Eq. (6-69),

$$\frac{d}{dt_2}\lambda_{mn}(t_2) = -\frac{i}{\hbar}\sum_k e^{(i/\hbar)(E_m-E_k)t_2}V_{mk}(t_2)\lambda_{kn}(t_2) - \frac{i}{\hbar}E_m\lambda_{mn}(t_2) \qquad (6\text{-}76)$$

Give a direct physical interpretation of this result. Next, deduce this result directly from the Schrödinger equation. *Hint:* Use Eq. (6-73) and substitute into the Schrödinger equation. Note that Eq. (6-76), with the initial condition $\lambda_{mn}(t_1) = \delta_{mn}$, could be used to determine the λ's directly.

We can interpret all of the terms in Eq. (6-69) by the rule that $-(i/\hbar)V_{mn}(t)\,dt$ is the amplitude that the potential V will scatter (or induce a transition) from a state n to a state m during the time interval dt. We can go from the state n to the state m by 0, 1, 2, . . . or more scatterings. We can go directly between the two states, i.e., with no scatterings, only if $m = n$. Thus the first term in the expansion is proportional to δ_{mn}.

The second term, given by Eq. (6-72), gives the amplitude that the transition will take place as the result of a single scattering. The amplitude for the particle to be found in the initial state n is $\exp[-iE_n(t_3 - t_1)/\hbar]$ at the time t_3. (In this case the phrase "to be found in the state n" should be interpreted as "available for scattering from the state n by the potential V.") The amplitude to be scattered by the potential V between the state n and the state m is $(-i/\hbar)V_{mn}$.

Finally, the amplitude to be found in the state m (which in this case means "the amplitude that the state m shall be available to the particle at the time the scattering takes place") at the time t_3 is proportional to $\exp[-iE_m(t_2 - t_3)/\hbar]$. This scattering can take place at any time between t_1 and t_2. Therefore, an integration over the time t_3 is carried out between these two end points.

The third term, given by Eq. (6-74), is the amplitude for a transition as the result of a double, or second-order, scattering. The first scattering takes the system from its initial state n to the intermediate state k at the time t_3. The system then stays in this state until the time t_4, when its availability for scattering is again measured by an exponential function, $\exp[(-i/\hbar)E_k(t_4 - t_3)]$. Another scattering takes place at the time t_4 and carries the system from the state k to the state n. We integrate over all of the possible alternate times for the scatterings t_4 and t_3, requiring only that the time t_3 shall be earlier than the time t_4. Next, we add over all the possible states k into which the system may have been scattered in the intermediate interval.

The terms of Eq. (6-69), which we have just interpreted, give the results of the general time-dependent perturbation theory. It is applicable when the unperturbed system has a *constant hamiltonian* and thus *definite energy values*. Next, we shall study some special cases of this theory in more detail.

First-order Transitions. First, let us take the case that the final state m is different from the initial state n and let us consider only the first Born approximation, i.e., the second term in Eq. (6-69). The result will be applicable for small values of V. The amplitude that we make the transition from m to n is

$$\lambda_{mn}{}^{(1)} = -\frac{i}{\hbar}\int_{t_1}^{t_2} e^{(i/\hbar)(E_n-E_m)t}V_{mn}(t)\,dt\,e^{-(i/\hbar)(E_nt_2-E_mt_1)} \tag{6-77}$$

This is a very important special formula in the time-dependent perturbation theory. Suppose as a first example that $V(x,t) = V(x)$ is not an explicit function of time. If we take the interval of time from 0 to T, then since V_{mn} is constant, we have

$$\begin{aligned}\lambda_{mn}{}^{(1)}e^{(i/\hbar)(E_nt_2-E_mt_1)} &= -\frac{i}{\hbar}V_{mn}\int_0^T e^{(i/\hbar)(E_n-E_m)t}\,dt\\ &= V_{mn}\frac{e^{(i/\hbar)(E_n-E_m)T}-1}{E_m-E_n}\end{aligned} \tag{6-78}$$

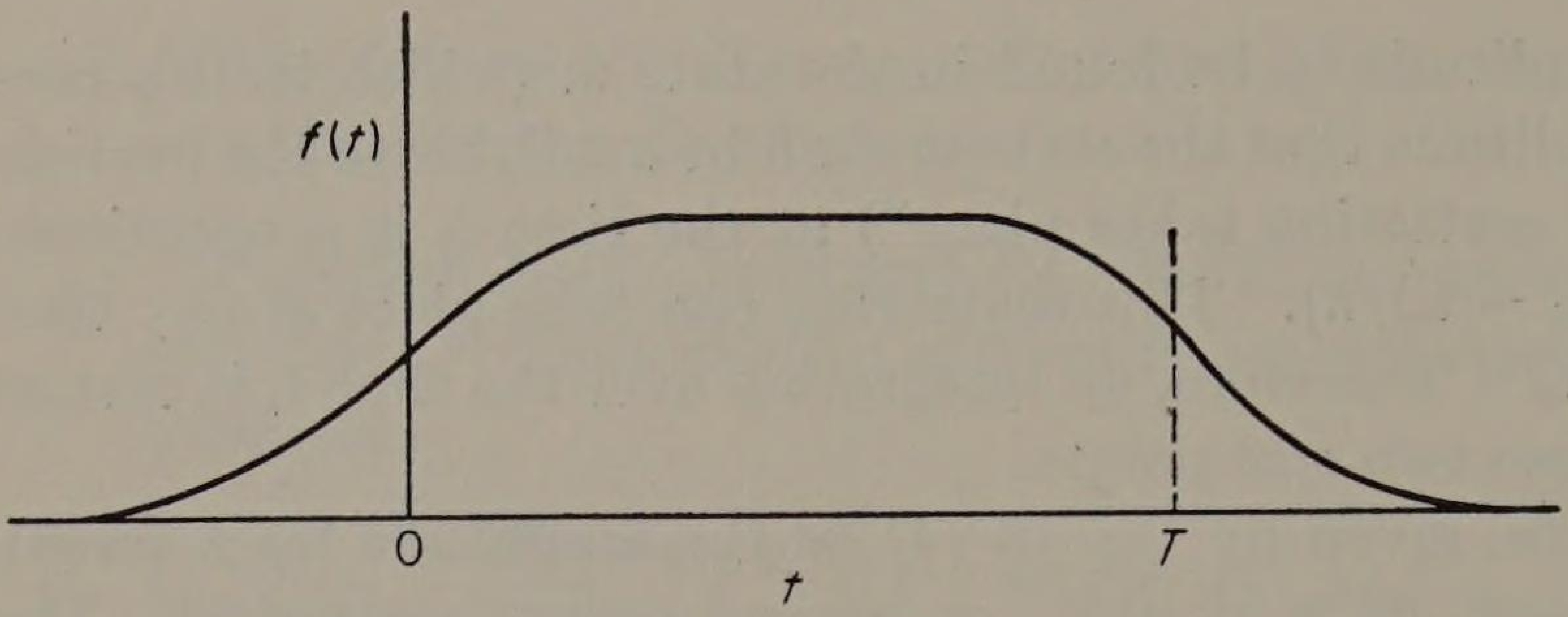


Fig. 6-12 The potential affecting the transition from m to n is turned on and off slowly with the time variation $f(t)$, shown here. As this time factor becomes smoother (e.g., as discontinuities appear in successively higher derivatives) the probability of a transition becomes smaller.

The probability of a transition during the time interval T is then

$$P(n \to m) = |\lambda_{mn}{}^{(1)}|^2 = |V_{mn}|^2 \left[4 \sin^2 \frac{(E_n - E_m)T}{2\hbar} \right] (E_n - E_m)^{-2} \tag{6-79}$$

We see that for at least a long interval T this probability is a rapidly oscillating function of the energy difference $E_n - E_m$. If E_n and E_m differ appreciably, it is very small, i.e., if $|V_{mn}| \ll |E_m - E_n|$. This means that the probability that the energy in the final state will be modified appreciably from that in the initial state by a very weak constant perturbation is very small. One might ask: How can the energy be expected to change at all by the large amount $E_m - E_n$ as a result of the small disturbance V_{mn}? The answer is that we have considered V to start suddenly at the time $t = 0$, and the definiteness of this time permits, by the uncertainty principle, a large uncertainty in the energy [cf. Eq. (5-19) and associated discussion].

Problem 6-20 Suppose V is turned on and off slowly. For example, let $V(x,t) = V(x)f(t)$, where $f(t)$ is smooth, as shown in Fig. 6-12.

$$\begin{aligned} f(t) &= \tfrac{1}{2}e^{\gamma t} && \text{for } t < 0 \\ &= 1 - \tfrac{1}{2}e^{-\gamma t} && \text{for } 0 < t < \frac{T}{2} \\ &= 1 - \tfrac{1}{2}e^{-\gamma(T-t)} && \text{for } \frac{T}{2} < t < T \\ &= \tfrac{1}{2}e^{-\gamma(t-T)} && \text{for } t > T \end{aligned} \tag{6-80}$$

Suppose that $1/\gamma$, which is the time of rise of the function $f(t)$, $\ll T$, but suppose further that $\gamma \ll (E_m - E_n)$. Show that the probability given by Eq. (6-79) is reduced by a factor $\{\gamma^2/[\gamma^2 + (E_m - E_n)^2]\}^2$. In this definition of $f(t)$ we still have a discontinuity in the second derivative with respect to time. Smoother functions make still further reductions.

If it should happen that E_m and E_n are exactly the same energy, we find $P(n \rightarrow m) = |V_{mn}|^2T^2/\hbar^2$. This grows as the square of the time. It means that a concept of the transition probability per unit time is not meaningful in this case. This formula holds only for T short enough that $V_{mn}T \ll \hbar$. It turns out that if only two states of exactly the same perturbed energy are involved, the probability of being found in the first goes as $\cos^2(|V_{mn}|T/\hbar)$ and of being found in the second as $\sin^2(|V_{mn}|T/\hbar)$, while our formula is only a first approximation to this.

Problem 6-21 Consider the special case that the perturbing potential V has no matrix elements except between the two levels 1 and 2; and further, suppose these levels are degenerate, that is, suppose $E_1 = E_2$. Let $V_{12} = V_{21} = v$ and let V_{11}, V_{22} and all other V_{mn} be 0. Show that

$$\lambda_{11} = 1 - \frac{v^2T^2}{2\hbar^2} + \frac{v^4T^4}{24\hbar^4} - \cdots = \cos\frac{vT}{\hbar} \tag{6-81}$$

$$\lambda_{12} = -i\frac{vT}{\hbar} + i\frac{v^3T^3}{6\hbar^3} - \cdots = -i\sin\frac{vT}{\hbar} \tag{6-82}$$

Problem 6-22 In Prob. 6-21 we have $V_{12} = V_{21}$, so that V_{12} is real. Show that even if V_{12} is complex, the physical results are the same (let $v = |V_{12}|$).

Such systems swing back and forth from one state to the other. A further conclusion can be drawn from this result. Suppose the perturbation acts for an extremely long time so that $V_{mn}T/\hbar \gg 1$. Then if the system is investigated at an arbitrary time T, which is somewhat indefinite, the probabilities of being in either the first or second state are, on the average, equal. That is, a small indefinite perturbation acting for a very long time between two states at the same energy makes these states have equal probability. This will be useful when we discuss the theory of statistical mechanics in Chap. 10.

The case of great importance is that in which the values allowed for E_m, the energy of the final state, are not separate and discrete but lie in a continuum, or at least are extremely closely spaced. Let us say that $\rho(E)\,dE$ is the number of levels or states in the range of energy E to $E + dE$. Then we can ask for the probability to go to some state in this continuum. First we see that to go to any state for which $E_n - E_m$ is large is very unlikely. It is most likely that the final state will be one of nearly the same energy as the original E_n (within an error $\pm V_{mn}$). The total chance to go into any state is

$$\sum_{m=1}^{\infty} P(n \rightarrow m) = \sum_{m=1}^{\infty} |V_{mn}|^2 \frac{4\sin^2[(E_m - E_n)T/2\hbar]}{(E_m - E_n)^2}$$

$$\simeq \int^{E_m} |V_{mn}|^2 \frac{4\sin^2[(E_m - E_n)T/2\hbar]}{(E_m - E_n)^2} \rho(E_m)\,dE_n \qquad (6\text{-}83)$$

The quantity $\{4\sin^2[(E_m - E_n)T/2\hbar]\}/(E_m - E_n)^2$ is very large if $E_m \approx E_n$, reaching a maximum of $T^2/\hbar^2$, whereas it is much smaller if E_m and E_n differ appreciably (relative to $\hbar/T$), as shown in Fig. 6-13. Thus almost all the contribution to the integral on E_m comes when E_m is in the neighborhood of the value E_n.

If V_{mn} does not vary rapidly so that we can replace it by a typical value, and furthermore if $\rho(E_m)$ likewise does not vary rapidly, then to a good approximation we can replace the integral of Eq. (6-83) by the expression

$$4|V_{mn}|^2\rho(E_n) \int^{E_m} \frac{\sin^2[(E_m - E_n)T/2\hbar]}{(E_m - E_n)^2}\,dE_m \qquad (6\text{-}84)$$

Since $\int_{-\infty}^{\infty} [(\sin^2 x)/x^2]\,dx = \pi$, the integral of Eq. (6-84) has the value $\pi T/2\hbar$ and we obtain the result that the probability for a transition to

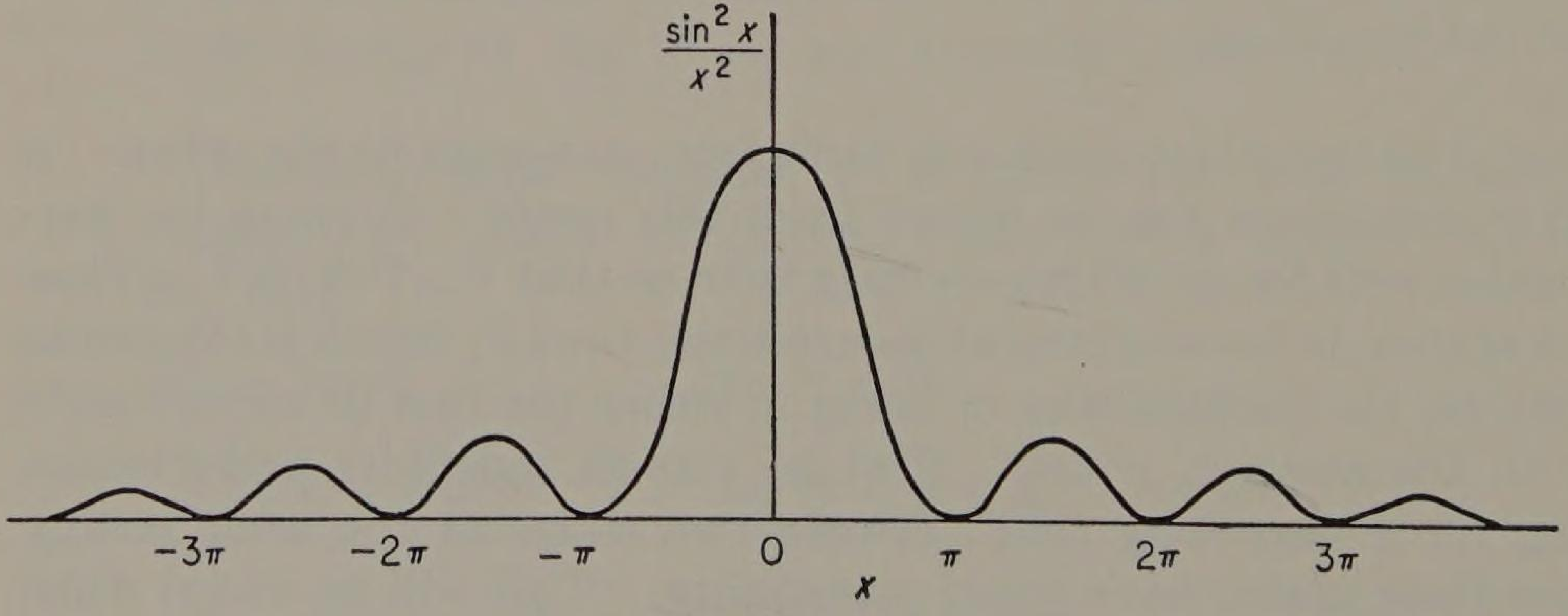


Fig. 6-13 In this figure the energy difference $E_m - E_n$ is replaced by the variable x. When these two energies are approximately equal (thus x is very small) the function $(\sin^2 x)/x^2$ approaches its maximum value. For large values of the difference the function becomes very small. Thus, in expressions involving this function, the most important contributions come from the central region, that is, the region where the two energies are approximately equal.

some state in the continuum is

$$P(n \rightarrow m) = 2\pi |V_{mn}|^2 \frac{\rho(E_n)T}{\hbar} \tag{6-85}$$

and that the energy in the final state is the same as the energy in the original state.

From these results we can write the probability of a transition per unit time in the form

$$\frac{dP(n \rightarrow m)}{dt} = \frac{2\pi}{\hbar} |M_{n\rightarrow m}|^2 \rho(E) \tag{6-86}$$

where $M_{n\rightarrow m}$ is called the *matrix element for the transition* and $\rho(E)$ is the density of levels in the final state. In our case $M_{n\rightarrow m}$ is V_{mn}. If we went to a higher-order expansion of λ_{mn}, it would be more complicated. Another way to write this expression is that the probability of a transition per unit time from state n to some particular state m is

$$\frac{dP(n \rightarrow m)}{dt} = \frac{2\pi \delta(E_n - E_m)|M_{n\rightarrow m}|^2}{\hbar} \tag{6-87}$$

Then when we sum over a group of states m, only those with $E_n = E_m$ survive. Since $\sum\limits_m \rightarrow \int dE_m\, \rho(E_m)$, we get as a result Eq. (6-86).

We may illustrate Eq. (6-86) by an example which we have previously discussed from a different point of view, namely, the scattering of an electron in a potential (cf. Sec. 6-4). Suppose an otherwise free particle has an interaction with a central potential $V(\mathbf{r})$ and we wish to discuss the scattering of this particle from an initial state of a definite momentum to final state in a new direction with another definite momentum. We suppose that the state n, the initial state, is a plane wave of momentum $\mathbf{p}_1$, so that the wave function ϕ_n is $\exp(i\mathbf{p}_1 \cdot \mathbf{r}/\hbar)$ (which is normalized so that the integral of the absolute square of ϕ_n over a unit volume is unity). Likewise, suppose the final state is a plane wave of momentum $\mathbf{p}_2$ so that the wave function ϕ_m is $\exp(i\mathbf{p}_2 \cdot \mathbf{r}/\hbar)$. The matrix element V_{mn} is

$$\begin{aligned} V_{mn} &= \int^{\mathbf{r}} e^{-(i/\hbar)\mathbf{p}_2\cdot\mathbf{r}} V(\mathbf{r}) e^{(i/\hbar)\mathbf{p}_1\cdot\mathbf{r}}\, d^3\mathbf{r} \\ &= v(\mathbf{q}) \end{aligned} \tag{6-88}$$

where $\mathbf{q} = \mathbf{p}_2 - \mathbf{p}_1$. In the scattering, the energy will be conserved so that $p_2^2/2m = p_1^2/2m$. This means that the magnitudes of momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ are the same. Let us set them equal to p, so that

$$|\mathbf{p}_2| = |\mathbf{p}_1| = p$$

By our usual convention for writing differential elements of momentum, the number of states which have their momenta in the volume element of momentum space $d^3\mathbf{p}_2$ is $d^3\mathbf{p}_2/(2\pi\hbar)^3 = p^2\,dp\,d\Omega/(2\pi\hbar)^3$, where $d\Omega$ is the element of solid angle which contains the momentum vector $\mathbf{p}_2$. An element dE of the energy range is connected to the element of momentum space by

$$dE = d\frac{p^2}{2m} = \frac{p\,dp}{m} \tag{6-89}$$

Thus the density of momentum states for particles traveling into the solid angle $d\Omega$ is

$$d\rho(E) = \frac{1}{dE}\frac{d^3\mathbf{p}_2}{(2\pi\hbar)^3} = \frac{mp\,d\Omega}{(2\pi\hbar)^3} = \rho(E)\,d\Omega \tag{6-90}$$

Substituting these relations into Eq. (6-86), we find the probability of transition per second into the element of solid angle $d\Omega$ is given by

$$\frac{dP}{dt} = \left(\frac{1}{2\pi\hbar^2}\right)^2 mp\,d\Omega\,|v(\mathbf{q})|^2 \tag{6-91}$$

We define an effective target area or cross section for scattering into $d\Omega$ as $d\sigma$ (cf. Sec. 6-4). The number of particles that will hit this area per second is the cross-sectional area times the velocity of the particles coming in, $u_1 = p_1/m$. This is because we have started with wave functions ϕ_n which were normalized on a unit volume; or in other words, the relative probability is unity to find one particle in any unit volume. Thus

$$\frac{dP}{dt}\,d\Omega = u_1\,d\sigma = \frac{p_1}{m}\,d\sigma \tag{6-92}$$

Therefore we obtain for the cross section the result

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 |v(\mathbf{q})|^2 \tag{6-93}$$

which is just what we obtained in Eq. (6-44).

Problem 6-23 Show that the same result is obtained for $d\sigma/d\Omega$ even if the wave functions ϕ_n are normalized to give a unit probability for some arbitrary volume V.

Problem 6-24 Suppose that the potential V is periodic in time. For example, suppose $V(x,t) = V(x)(e^{i\omega t} + e^{-i\omega t})$. Show that the probability for a transition to take place is small unless the final state is one of the two values (1) $E_{\text{final}} = E_{\text{initial}} + \hbar\omega$ (corresponding to an absorption of energy) or (2) $E_{\text{final}} = E_{\text{initial}} - \hbar\omega$ (corresponding to the emission of energy). This means that Eq. (6-86) is unchanged, but the density $\rho(E)$ of states must be calculated at these new values of E. Or, in analogy with Eq. (6-87), we have

$$\frac{dP(n \rightarrow m)}{dt} = \frac{2\pi}{\hbar} |M_{n\rightarrow m}|^2[\delta(E_m - E_n - \hbar\omega) + \delta(E_m - E_n + \hbar\omega)] \tag{6-94}$$

Problem 6-25 It has been argued that the equations of the electrodynamics must, like those of mechanics, be converted to a quantized form on the basis of the photoelectric effect. Here an electron of energy $\hbar\omega$ is occasionally emitted from a thin layer of metal under the influence of light of frequency ω. Is this impossible if matter obeys the quantum laws but light is still represented as a continuous wave? What arguments can you adduce for the necessity of giving up a classical description of electrodynamics, in view of the results of Prob. 6-24?

Problem 6-26 Suppose we have two discrete energy levels E_1 and E_2, neither of which is in the continuum. Let a transition be induced by a potential of the form $V(x,t) = V(x)f(t)$. Show that the probability of transition is

$$P(1 \rightarrow 2) = |V_{12}|^2|\phi(\omega_0)|^2 \tag{6-95}$$

if $f(t)$ is representable by the Fourier transform

$$f(t) = \int_{-\infty}^{\infty} \phi(\omega)e^{i\omega t}\,\frac{d\omega}{d\pi} \tag{6-96}$$

and $\omega_0 = (E_2 - E_1)/\hbar$.

If $f(t)$ is a statistically irregular function familiar from the theory of noise (called filtered white noise), the value of $\phi(\omega)$ given by the inverse transform

$$\phi(\omega) = \int_{-T}^{T} f(t)e^{-i\omega t}\,dt \tag{6-97}$$

depends on the range T of integration of the time t. If T is very large, $|\phi(\omega_0)|^2$ can be shown to be proportional to T. Thus we get a transition probability proportional to the time and the "intensity" or "power" (mean-square value of f per second) at frequency ω_0 per unit frequency range. In virtue of this, the probability for the transition of an atom in a continuous spectrum of light is proportional to (1) the exposure time and (2) the intensity of light at the frequency $(E_2 - E_1)/\hbar$ of absorption.

The Higher-order Terms. It is interesting to look at the second-order term in the perturbation expansion. This term is of special importance in the problems where $V_{mn} = 0$ for those particular states m and n of interest. Let us suppose that we have such a problem, and suppose further that there are other states $k \neq m$ for which $V_{km} \neq 0$. The first-order term is 0, and so long as $n \neq m$ the zero-order term is likewise zero. Thus the lowest-order term which enters into the calculation of the transition amplitude is the second.

Suppose that the potential V is independent of t. Then the second-order term in the transition element is $\lambda_{mn}{}^{(2)}$; and if $T = t_2 - t_1$, we have from Eq. (6-74)

$$\begin{aligned}
e^{+(i/\hbar)(E_m t_2 - E_n t_1)} \lambda_{mn}{}^{(2)} &= -\frac{1}{\hbar^2} \sum_k V_{mk} V_{kn} \int_0^T dt_4 \int_0^{t_4} dt_3 \\
&\qquad \times e^{(i/\hbar)(E_m - E_k)t_4} e^{(i/\hbar)(E_k - E_n)t_3} \\
&= \frac{i}{\hbar} \sum_k V_{mk} V_{kn} \int_0^T e^{(i/\hbar)(E_m - E_k)t_4} \big(e^{(i/\hbar)(E_k - E_n)t_4} - 1\big) \\
&\qquad \times \frac{dt_4}{E_k - E_n} \\
&= \sum_k \frac{V_{mk} V_{kn}}{E_k - E_n} \left(\frac{e^{(i/\hbar)(E_m - E_n)T} - 1}{E_m - E_n} \right. \\
&\qquad \left. - \frac{e^{(i/\hbar)(E_m - E_k)T} - 1}{E_m - E_k} \right)
\end{aligned} \tag{6-98}$$

The first of the two terms in the last factor of this result has the same time dependence as we have seen in our first-order result. Therefore if the other term is neglected for a moment, we see that the net result would again be to make transitions to states where $E_m = E_n$, with a probability proportional to T. The probability per unit time is of the same form as Eq. (6-86) but with $M_{n \to m}$ now given by

$$M_{n \to m} = \sum_k \frac{V_{mk} V_{kn}}{E_k - E_n} \tag{6-99}$$

If we assume that the states lie in a continuum, the sum becomes an integral. Equation (6-99) is correct in the circumstance that it is impossible to go by first-order transition, impossible not only to the state m but also to any state k of the same energy as the initial state.

Under these circumstances $V_{kn} = 0$ for states such that $E_k = E_n$. Then the second term in Eq. (6-98) is never large; for it cannot be large unless $E_n - E_k$ is nearly zero, and then V_{kn} in the numerator is zero. All the effects come from the first term, and Eq. (6-99) is correct. Furthermore, in the sum over k in Eq. (6-98) there is no ambiguity at the pole where $E_k = E_m$; for the numerator vanishes at this same value of E_k.

On the other hand, in some situations it may be true that a first-order transition is possible to some other continuum state (e.g., a nucleus may decay in more than one way). In such a case the sum in Eq. (6-99) is meaningless; for we must define what to do near the pole. It is the neglected second term in Eq. (6-98) which comes to our rescue here and shows that the correct expression for $M_{n\to m}$ (now including the first-order term also for generality) is

$$M_{n\to m} = V_{mn} + \sum_k \frac{V_{mk}V_{kn}}{E_k - E_n - i\epsilon} \tag{6-100}$$

in the limit $\epsilon \to 0$. How this comes about we shall now analyze.

First we may notice that for large T we cannot get a large probability of transition (proportional to T, that is) unless E_n and E_m are practically equal (within about $\hbar/T$). This is evident for the first term in Eq. (6-98). For the second term large amplitudes can arise only if $E_k \approx E_m$; but if E_m is not very close to E_n, the factor in front is a smooth function of E_k for E_k near E_m. Taking it as nearly constant for a small range near $E_k = E_m$, we see that the second term can be approximated by some constant times

$$\frac{e^{(i/\hbar)\epsilon T} - 1}{\epsilon}\, d\epsilon$$

where $\epsilon = (E_m - E_k)$ is to be integrated over a small range, say $-\delta$ to $+\delta$. But

$$\int_{-\delta}^{\delta} \frac{e^{(i/\hbar)\epsilon T} - 1}{\epsilon}\, d\epsilon = \int_{\delta T/\hbar}^{\delta T/\hbar} \frac{e^{iy} - 1}{y}\, dy = \int_{\delta T/\hbar}^{\delta T/\hbar} \left(\frac{\cos y - 1}{y} + \frac{i \sin y}{y}\right) dy \tag{6-101}$$

The first integral is that of an odd function and vanishes. The second approaches a finite limit as $T \to \infty$ (and therefore as $\delta T/\hbar \to \infty$). That is,

$$2i \int_0^\infty \frac{\sin y}{y}\, dy = 2\pi i$$

so no large transition probability occurs. A large effect can arise only in case E_n and E_m are essentially equal, for then the double coincidence of the two poles from $(E_k - E_n)^{-1}$ and $(E_m - E_k)^{-1}$ can make the second term important. Therefore, we continue the analysis, assuming E_m and E_n are nearly equal.

The sum of k in Eq. (6-98) can be divided into two regions by choosing a very small energy Δ and breaking the sum up into a part A for which $|E_k - E_n| \geq \Delta$ and a part B for which $|E_k - E_n| < \Delta$. We choose Δ to be small enough that the factor $V_{mk}V_{kn}$ does not vary appreciably when E_k varies around E_n over this energy range 2Δ. This is some finite energy, and we shall take T so long that $\hbar/T \ll \Delta$, which means that $|E_n - E_m| \ll \Delta$.

First for part A, $|E_k - E_m| \geq \Delta$. Then the second term cannot become large; for its poles are avoided. Only the first contributes, and the contribution is

$$a\frac{e^{ix} - 1}{x}\frac{T}{\hbar} \tag{6-102}$$

where $x = (E_m - E_n)T/\hbar$ and

$$a = \sum_k^{(A)} \frac{V_{mk}V_{kn}}{E_k - E_n}$$

The sum extends over all E_k except for within $\pm\Delta$ of E_m. This sum is nearly independent of Δ, and as $\Delta \to 0$ it is the definition of a principal-value integral. That is, in the limit $\Delta \to 0$ we can write

$$a = V_{mn} + \sum_k V_{mk}V_{kn}\,\text{P.P.}\,\frac{1}{E_k - E_n} \tag{6-103}$$

where P.P. is the principal part and we have reinstated the first-order term, in case it does not vanish.

For the region B we take $V_{mk}V_{kn}$ to be constant at its value for $E_k - E_m = 0$. That is, we replace $\sum_k^{(B)} V_{mk}V_{kn}F(E_k)$ by

$$\left[\sum_k V_{mk}V_{kn}\delta(E_k - E_m)\right] \int_{E_m-\Delta}^{E_m+\Delta} F(E_k)\,dE_k \tag{6-104}$$

We write this as bI, where

$$b = \sum_k V_{mk}V_{kn}\delta(E_k - E_m) \tag{6-105}$$

and

$$I = \int_{E_m-\Delta}^{E_m+\Delta} \frac{dE_k}{E_k - E_n}\left(\frac{e^{(i/\hbar)(E_m-E_n)T} - 1}{E_m - E_n} - \frac{e^{(i/\hbar)(E_m-E_k)T} - 1}{E_m - E_k}\right) \tag{6-106}$$

Now we put $(E_m - E_n)(T/\hbar) = x$ and $(E_k - E_n)(T/\hbar) = y$, so that $(E_m - E_k)(T/\hbar) = x - y$, to get

$$I = \frac{T}{\hbar}\int_{-\Delta T/\hbar}^{\Delta T/\hbar} \frac{dy}{y}\left(\frac{e^{ix} - 1}{x} - \frac{e^{i(x-y)} - 1}{x - y}\right) \tag{6-107}$$

This integral is most easily evaluated by contour integration, imagining y as a complex variable and changing the contour. Instead of integrating on the straight line from $-\Delta T/\hbar$ to $\Delta T/\hbar$, we go on the semicircle of radius $\Delta T/\hbar$ below the real axis. Since $\Delta T/\hbar$ is very large, the second term contributes negligibly; and since

$$\int_{-\Delta T/\hbar}^{\Delta T/\hbar} \frac{dy}{y} = i\pi$$

on this contour, we get $I = i\pi(T/\hbar)(e^{ix} - 1)/x$. Putting the A and B parts together, we get

$$(a + i\pi b)\,\frac{(e^{ix} - 1)T}{x\hbar} \tag{6-108}$$

for the amplitude. This gives a probability for transition of the form Eq. (6-86) with

$$M_{n\to m} = a + i\pi b = V_{mn} + \sum_k V_{mk}V_{kn}\left[\text{P.P.}\,\frac{1}{E_k - E_n} + i\pi(E_k - E_m)\right] \tag{6-109}$$

The last bracket can be written $(E_k - E_m - i\epsilon)^{-1}$ in the limit as $\epsilon \to 0$, as we have written in Eq. (6-100).

From Eq. (6-100) we learn then that even if no direct transition is possible from $n \to m$, nevertheless the transition can occur, as we say, through a *virtual* state. That is, we can imagine that the system goes from n to k, then k to m. The amplitude for an indirect transition process is given by Eq. (6-99). We note that it is not right to say that it actually goes through one or another intermediate state k, but rather that in characteristic quantum-mechanical fashion there is a certain amplitude to go via the various intermediate states k, and the contributions interfere.

The intermediate states are not of the same energy as the initial and final states. The conservation of energy is not violated, for the virtual state is not permanently occupied. The strength of contribution to the sum varies inversely with this energy discrepancy.

There is nothing absolute about these intermediate states. They come from considering V as a perturbation to a system H and from speaking about the true states of $H + V$ in terms of those of H alone. If other separations are made as to what is the "unperturbed" problem and what is the "perturbation," different formulas and intermediate states will arise in the description.

When the potential depends upon time (e.g., periodically), many interesting effects result. Most of these have been observed in microwave experiments, where the perturbation $V(x,t)$ is a weak electric or magnetic field with a periodic variation in time.

Problem 6-27 Derive the perturbation expansion up through the terms of the second order for potentials periodic in time.

Sometimes a transition cannot take place except by the use of two or more intermediate virtual states. Analysis of such transitions requires the calculation of third- and higher-order terms in the perturbation expansion.

Problem 6-28 Show that when a transition is impossible either directly or through a single intermediate state, but requires the use of two intermediate states, it is determined by the matrix element given by

$$M_{n\to m} = \sum_k \sum_l \frac{V_{mk} V_{kl} V_{ln}}{(E_m - E_k)(E_m - E_l)} \tag{6-110}$$

This corresponds to the third-order term in the perturbation expansion.

Problem 6-29 Suppose two perturbations are acting, $V(x,t)$ and $U(x,t)$, for example, a combination of d-c and a-c electric fields or a combination of electric and magnetic fields. Suppose further that a certain transition cannot occur with either V or U alone, but can occur only when both act together. Under the special assumption that both V and U are constant in time, show that the matrix element determining the transition element is given by

$$M_{m\to n} = \sum_k \frac{V_{mk} U_{kn} + U_{mk} V_{kn}}{E_m - E_k} \tag{6-111}$$

Next, suppose both potentials are periodic in time but have different frequencies, ω_1 and ω_2. What then is the matrix element?

Calculation of the Change in Energy of the State. In computing transition amplitudes we have considered only those states $n \neq m$. Suppose we turn our attention to the term $m = n$. Considering the zero- and first-order terms in the perturbation expansion, we have

$$\lambda_{mm} = 1 - \frac{i}{\hbar} \int_0^T V_{mm}(t)\, dt \tag{6-112}$$

If V is constant in time, this gives $1 - (i/\hbar)V_{mm}T$. What is the meaning of this result? As a consequence of the introduction of the additional potential V into the original hamiltonian we can expect the energies of all of the states of the system to be slightly altered. We can write the new energy of the state m as $E_m + \Delta E_m$. The time-dependent portion of the wave function describing this state will be $\exp\,[(-i/\hbar)(E_m + \Delta E_m)t]$ instead of the previous function $\exp\,(-i/\hbar)E_m t$.

Over the period of time T during which the perturbing potential acts this relative difference in phase introduces the factor

$$\exp\,[(-i/\hbar)\,\Delta E_m\, T]$$

Expanding this factor to first order in time gives $1 - (i/\hbar)\,\Delta E_m\, T$. Thus we see that a first-order calculation of the energy shift in a state m due to a perturbation V is

$$\Delta E_m = V_{mm} \tag{6-113}$$

This derivation of the first-order energy shift is not satisfactory if the system is degenerate, i.e., if there are initially very many states of exactly the same energy. It turns out that in such a case terms of second order in V give equally large effects.

Adding in the second-order term in the perturbation expansion for the transition element gives

$$e^{-(i/\hbar)E_mT}\lambda_{mm} = 1 - \frac{i}{\hbar} V_{mm}T - \left(\frac{1}{\hbar}\right)^2 \sum_k \int_0^T \int_0^{t_4} e^{-(i/\hbar)(E_k - E_m)(t_4 - t_3)}\, dt_3\, dt_4\, V_{mk}\, V_{km} \tag{6-114}$$

For the present let us assume that there is no degeneracy. Consider first the term $k = m$ in the series which is the second-order term. The integral over this particular term is just $T^2/2$. Integrals for the terms $k \neq m$ can also be performed easily to give the result

$$\begin{aligned} e^{-(i/\hbar)E_mT}\lambda_{mm} &= 1 - \frac{i}{\hbar}V_{mm}T - \frac{1}{2\hbar^2}V_{mm}{}^2T^2 \\ &\quad - \sum_{k\neq m}\frac{i|V_{km}|^2}{(E_m - E_k)\hbar}\left\{T - \frac{1 - \exp\left[-iT(E_k - E_m)/\hbar\right]}{(i/\hbar)(E_k - E_m)}\right\} \end{aligned} \tag{6-115}$$

The first three terms on the right-hand side of this equation represent an expansion through second order of $\exp(-iV_{mm}T/\hbar)$. The first of the summation terms, the one proportional to T, can be interpreted as a second-order energy change. That is, the incremental energy is not just V_{mm}, but contains higher-order corrections. Writing out the energy correction through second order in the perturbation energy, we get

$$\Delta E_m = V_{mm} - \sum_{k\neq m}\frac{V_{mk}V_{km}}{E_m - E_k} \tag{6-116}$$

This last equation gives the correct expression, through second order, for the shift in energy of nondegenerate states. This result is much more easily obtained by conventional methods, i.e., by finding solutions of

$$(H + V)\phi = E\phi \tag{6-117}$$

Furthermore, the conventional approach based on Eq. (6-117) permits simpler handling of degenerate states. However, it has been our purpose here to give an example of the use of transition amplitudes, rather than to give the simplest formulas for the computation of energy shifts.

Actually, there are more complex problems involving energy increments in which the method of transition amplitudes is the simplest to apply. In such applications the scheme, as we have attempted to show above, is to identify terms in a series proportional to T, T^2, etc. Then, if we remember that the amplitude to stay in the initial state is proportional to $\exp(-i\,\Delta E\;T/\hbar)$ and that the series expansion is equivalent to a series expansion of this exponential, the correct expression for ΔE can be written down.

We have still not discussed the last term in Eq. (6-115). If the states E_k lie in a continuum, we must also define the character of the reciprocal in the sum of Eq. (6-116). If we take it to mean the principal value, just as we found when analyzing the problem in second order for $n \neq m$, this extra term can be shown to produce an effect proportional to T and to lead to an additional correction to Eq. (6-116) of

$$\Delta' E_m = -i\pi \sum_k \delta(E_m - E_k) V_{mk} V_{km} \tag{6-118}$$

But this cannot represent a further correction to the energy for it is purely imaginary, and the energy must be real. Let us call it $-i\gamma/2$ (the ½ is for convenience later) and write

$$\Delta E_m - \frac{i\gamma}{2} = V_{mm} - \sum_k \frac{|V_{mk}|^2}{E_m - E_k - i\epsilon} \tag{6-119}$$

This implies that the transition amplitude λ_{mm} to be in the mth state after a long time is proportional to

$$\exp\left[-i\left(\Delta E_m - \frac{i\gamma}{2}\right)T\right] = \exp\left[-i(\Delta E_m)T\right] \exp\left(-\frac{\gamma T}{2}\right)$$

The first factor is the energy shift. The second is easily interpreted; for the probability to be in state m after time T is $|\lambda_{mm}|^2 = \exp(-\gamma T)$. It falls with time because at each instant there is a probability that a transition is made from m to some other state. That is, if all is consistent, γ must be the total probability per second of a transition from m to any state in the continuum of the same energy. This it is, because from Eq. (6-118) our γ is

$$\gamma = \sum_k 2\pi\delta(E_m - E_k)|V_{mk}|^2 \tag{6-120}$$

So we see that the total probability per second is just the sum of Eq. (6-87) over all possible final states as required (i.e., up to the required order in V).

The reciprocal of γ is called the mean lifetime of the state. Strictly speaking, a state with a finite lifetime has no definite energy; the energy uncertainty by the Heisenberg relation is $\hbar$/lifetime, or γ.

If resonance experiments are performed to find the energy difference of two levels, each of which has a decay rate γ, the resonance is not sharp but has a definite shape. The center of the resonance determines the energy difference, and the width of the resonance gives the sum of the γ's of each level.

7

Transition Elements

In the preceding chapter we developed the concept of a perturbation treatment for changes of state in a quantum-mechanical system. We carried out an investigation of this method as it is applied to systems whose unperturbed hamiltonians are constant in time. In this chapter we shall continue the development of the perturbation concept and generalize the treatment to cover systems where the unperturbed state may have a hamiltonian varying with time. We shall introduce a more general type of notation and attempt to broaden and deepen our understanding of the ways in which changes of state take place in a quantum-mechanical system. The notation to be introduced applies to a type of function which will be defined in the first portion of this chapter. The function is called a *transition element.*

The chapter is divided into four parts. The first part, comprising Sec. 7-1, gives a definition of transition amplitudes and transition elements, with the help of examples based upon the perturbation theory of Chap. 6. The second part, comprising Secs. 7-2 to 7-4, gives some interesting general relations among transition elements. The third part, consisting of Sec. 7-5, shows the connection between transition elements defined with the help of path integrals and the treatment of quantum-mechanical transitions defined in terms of the more usual operator notation of quantum mechanics. In the last part, comprising Secs. 7-6 and 7-7, the results learned in the preceding sections are applied to two interesting problems of quantum mechanics.

7-1 DEFINITION OF THE TRANSITION ELEMENT

The development of a quantum-mechanical system with time can be pictured as follows. At an initial time t_1 the state is described by the wave function $\psi(x_1,t_1)$. At a later time t_2 this original state will develop into the state $\phi(x_2,t_2)$.

At this later time suppose we ask the question: What is the probability of finding the system in the specific state $\chi(x_2,t_2)$? We know from the general principles developed in Chap. 5 that the probability of finding the system in this specified state is proportional to the square of the amplitude defined by

$$\int \chi^*(x_2,t_2)\phi(x_2,t_2)\,dx_2$$

We also know from Chap. 3 that the function ϕ can be expressed in terms of the original wave function with the help of the kernel K describing the propagation of the system between the times t_1 and t_2.

Thus, in determining the probability of finding the system in a specified state we can start with the original wave function ψ and bridge the time gap with the propagation kernel $K(2,1)$.

The resulting amplitude, whose absolute square gives the probability desired, we shall call the *transition amplitude*, and we shall write it in the following notation:

$$\langle\chi|1|\psi\rangle = \int\int \chi^*(x_2)K(2,1)\psi(x_1)\,dx_2\,dx_1 \tag{7-1}$$

We wish to return to an even more basic description of the transition phenomena, and we reintroduce the action S describing the behavior of the system between two time limits. Thus we write the transition amplitude as

$$\langle\chi|1|\psi\rangle_S = \int\int\int_{x_1}^{x_2} \chi^*(x_0)e^{iS/\hbar}\psi(x_1)\,\mathfrak{D}x(t)\,dx_1\,dx_2 \tag{7-2}$$

Here we have made the notation a bit more explicit by attaching subscript S to the transition amplitude to indicate the action for which the integral was calculated. The path integral is to be taken over all paths which go from x_1 to x_2, and the result of this path integral is multiplied by the two wave functions, then integrated over the space variables at the two limits.

Before proceeding further, we shall define the notation more completely to cover a more general situation. We introduce the functional $F[x(t)]$ without (for the present) describing its physical nature. With this functional we define a *transition element* as follows:

$$\langle\chi|F|\psi\rangle_S = \int\int\int \chi^*(x_2)F[x(t)]e^{iS/\hbar}\psi(x_1)\,\mathfrak{D}x(t)\,dx_1\,dx_2 \tag{7-3}$$

Here F is any functional of $x(t)$ which does not involve $x(t)$ at the end points x_1 or x_2 or beyond the end points. In the special case that $F = 1$, the integral of Eq. (7-3) defines a transition amplitude.

It is difficult to understand transition elements on the level of intuitive physics. One approach toward such understanding involves a classical analogy. Picture a small particle moving with brownian motion. At some initial time, $t = t_1$, the particle is at x_1. We wish to determine the probability that the particle arrives at the point x_2 at the time $t = t_2$. For quantum-mechanical particles, we talk about starting from an initial state and arrive at some final state. Thus, the point x_1 for the brownian particle is analogous to the initial wave function $\psi(x_1)$ in Eq. (7-2), and the point x_2 to $\chi(x_2)$. Furthermore, the solution of the quantum-mechanical problem requires integration over the variables x_1 and x_2 of the initial and final states—a step unnecessary in our classical problem.

We would solve the classical problem by considering all possible paths for the particle's motion. We would weight each path with the function defining the probability that the particle actually follows such a path and then integrate the weighted contributions for all such paths. The weighting function is analogous to the term $e^{iS/\hbar}$ appearing in the integral of Eq. (7-2).

The final position in such a problem would be not a single point, but rather a small region, x_2 to $x_2 + dx$. The result, when properly normalized, would be the distribution function $P(x_2)$ giving the relative probability of arriving in the (differential) vicinity of x_2. This function is analogous to the transition amplitude of Eq. (7-2) in the case that ψ and χ are δ functions of position.

Now suppose we wish to know more about the motion than simply the relative probability to arrive at x_2. For example, we may wish to find the acceleration experienced by the particle at some particular time, say 1 sec after it starts. But now we need the weighted average of the acceleration, i.e., the acceleration for each possible path with each path weighted by the function defining the probability of the path. Such a weighted average is analogous to the transition element of Eq. (7-3). The property of interest, such as the acceleration at some time t, replaces the function $F[x(t)]$ in the integral of Eq. (7-3). The classical problem could be solved by a path integral very similar in form to that of Eq. (7-3).

In the remainder of this chapter we shall make use of this analogy, and we shall occasionally refer to transition elements as "weighted averages." However, it must be kept in mind that the weighting function in quantum mechanics is a complex function. Thus the result is not an "average" in the ordinary sense.

The path integral method of solving brownian-motion problems as described in this classical analogy is actually a very powerful method. It will be developed in detail in Chap. 12. For now, we attempt to further clarify the notion of a transition element with the help of the perturbation theory developed in Chap. 6.

Perturbations. Suppose the action describing the development of the system can be separated into two parts, so that $S = S_0 + \sigma$. We suppose that the first part S_0 leads to simple path integrals, whereas the remaining part σ is small enough that we can apply a perturbation scheme. We write the exponential function of Eq. (7-2) as

$$e^{iS/\hbar} = e^{iS_0/\hbar}e^{i\sigma/\hbar} \tag{7-4}$$

Using Eq. (7-3), the transition element of Eq. (7-2) becomes

$$\langle \chi|1|\psi\rangle_{S_0+\sigma} = \langle \chi|e^{i\sigma/\hbar}|\psi\rangle_{S_0} \tag{7-5}$$

The exponential function can be expanded to give

$$\langle \chi|1|\psi\rangle_{S_0+\sigma} = \langle \chi|1|\psi\rangle_{S_0} + \frac{i}{\hbar}\langle \chi|\sigma|\psi\rangle_{S_0} - \frac{1}{2\hbar^2}\langle \chi|\sigma^2|\psi\rangle_{S_0} + \cdots \tag{7-6}$$

This expansion is a generalized version of Eq. (6-3) and forms the basis of the perturbation theory. The transition elements which arise in most quantum-mechanical problems result from this expansion.

Suppose the perturbation action σ results from a perturbation potential, so that

$$\sigma = \int V[x(t),t]\,dt \tag{7-7}$$

Then the first-order perturbation is given by the transition element

$$\langle \chi|\sigma|\psi\rangle_{S_0} = \int \langle \chi|V[x(t),t]|\psi\rangle_{S_0}\,dt \tag{7-8}$$

To evaluate this element, we need to solve the integral

$$\langle \chi|V[x(t),t]|\psi\rangle_{S_0} = \iiint_{x_1}^{x_2} \chi^*(x_2)e^{iS_0/\hbar}V[x(t),t]\psi(x_1)\,dx_1\,dx_2\,\mathcal{D}x(t) \tag{7-9}$$

The first step in the solution of this integral is the same as the solution for the perturbation kernel $K^{(1)}$ described in Eqs. (6-8) to (6-11). This solution for the path integral is followed by integration over both end points, x_1 and x_2, as well as an integral over the midpoint x_3 [identified as the point c in Eq. (6-10)]. That is,

$$\langle \chi|V[x(t),t]|\psi\rangle_{S_0} = \int \chi^*(x_2)K_0(2,3)V(3)K_0(3,1)\psi(x_1)\,dx_1\,dx_2\,dx_3 \tag{7-10}$$

We have now arrived at an expression which combines three concepts previously introduced. First, we have made use of the propagation rule for a wave function as defined in Eq. (3-42). Next, we have made use of the amplitude function as defined in Eq. (5-31), which gives the amplitude that a system known to be in one state will be found in another state. Lastly, we have made use of the first-order perturbation theory given in Eq. (6-11) for the kernel describing the propagation in time. All of these ideas combined give the transition element of Eq. (7-10). The absolute square of this element is the probability that a system starting in state ψ and acted upon by the small potential $V(x,t)$ will be found at a later time in the state χ (if state χ would not be reached for $V = 0$, that is, if $\langle \chi|1|\psi\rangle = 0$).

We can use Eq. (3-42) to shorten our notation, in the same way that the notation of Eq. (6-23) was shortened into the form of Eq. (6-25). We defined function $\psi(x_3,t_3)$ as

$$\psi(3) = \int K_0(3,1)\psi(x_1)\,dx_1 \tag{7-11}$$

which is the wave function that would result at the time t_3 from the initial wave function if there were no perturbation acting. In a similar way we define

$$\chi^*(x_3,t_3) = \int \chi^*(x_2)K_0(2,3)\,dx_2 \tag{7-12}$$

as the complex conjugate of the wave function which at t_3 would lead to the function $\chi(x_2)$ at the time t_2 if there were no perturbation acting. [See Eq. (4-38) and the following discussion, including Prob. 4-7.]

In terms of these new wave functions, the first-order term in the perturbation expansion can be simplified to read

$$\langle\chi|\int V[x(t),t]\,dt|\psi\rangle_{S_0} = \int\int \chi^*(3)V(3)\psi(3)\,dx_3\,dt_3 \tag{7-13}$$

We see here that the transition amplitude written in this form is a generalization of the transition amplitude λ_{mn} which was introduced in Sec. 6-5. If the wave functions on the right-hand side of Eq. (7-13) are eigenfunctions, then the resulting transition amplitude is identical with $\lambda_{mn}^{(1)}$, as defined by Eq. (6-70).

Thus the evaluation of a transition element of a functional $F[x(t)]$, which depends only on x at a particular time t [that is, an ordinary function of $x(t)$], or of a time integral of such a functional presents no problem. The evaluation of a transition element for functionals involving the values of x at two separate times is also easy. This occurs, for example, in the second-order perturbation term. This can be written as

$$\frac{1}{2\hbar^2}\langle\chi|\sigma^2|\psi\rangle_{S_0} = \frac{1}{2\hbar^2}\int\int \langle\chi|V[x(t),t]V[x(s),x]|\psi\rangle\,dt\,ds \tag{7-14}$$

The integrand of this last equation is itself a transition element, and it is written as

$$\langle\chi|V[x(t),t]V[x(s),s]|\psi\rangle = \int\int \chi^*(4)V(4)K_0(4,3)V(3)\psi(3)\,dx_3\,dx_4 \tag{7-15}$$

where we have substituted $t_3 = s$ and $t_4 = t$ if $s < t$ or $t_3 = t$ and $t_4 = s$ if $s > t$.

Thus the second-order term in the perturbation expansion becomes

$$\frac{1}{2\hbar^2}\left\langle \chi \left| \int V[x(t),t]\,dt \int V[x(s),s]\,ds \right| \psi \right\rangle = \iint \chi^*(4)V(4)K_0(4,3)\psi(3)\,dx_3\,dt_3\,dx_4\,dt_4 \quad (7\text{-}16)$$

This can be recognized as a generalization of the transition amplitude defined in Eq. (6-74). Expressions involving three or more functions are also readily written down.

Equation (7-4) corresponds also to a more general type of perturbation theory. For example, consider the case of the particle interacting with an oscillator. After integrals have been carried out over the coordinates describing the oscillator, the resulting action can be written as $S_0 + \sigma$, where (see Sec. 3-10)

$$\sigma = \frac{1}{m\omega \sin \omega T} \int_{t_1}^{t_2} \int_{t_1}^{t} g[x(t),t]g[x(s),s] \sin \omega(t_2 - t) \sin \omega(s - t_1)\,ds\,dt \quad (7\text{-}17)$$

with $g[x(t),t]$ characterizing the interaction of the particle and oscillator, and $T = t_2 - t_1$.

We have noted that path integrals involving such complicated actions are very hard to evaluate indeed; but if the effect of the complicated term σ is expected to be small, we can obtain useful results with less effort with the help of the perturbation expansion of Eq. (7-4). To illustrate, we find the first-order term in such an expansion (i.e., the first Born approximation). Using Eq. (7-17) for σ, we must evaluate the term $(i/\hbar)\langle\chi|\sigma|\psi\rangle_{S_0}$. This term can be written as

$$\frac{i}{\hbar}\langle\chi|\sigma|\psi\rangle_{S_0} = \frac{i}{\hbar m\omega \sin \omega T} \int_{t_1}^{t_2} \int_{t_1}^{t} \langle|g[x(t),t]g[x(s),s]|\psi\rangle_{S_0} \sin \omega(t_2 - t) \times \sin \omega(s - t_1)\,ds\,dt \quad (7\text{-}18)$$

so that the difficult part of the problem is reduced to finding

$$\langle\chi|g[x(t),t]g[x(s),s]|\psi\rangle_{S_0}$$

But this we have already done in Eq. (7-15), except that g replaces V. Therefore, we write

$$\langle\chi|g[x(t),t]g[x(s),s]|\psi\rangle_{S_0} = \iint \chi^*(4)g[x(t_4),t_4]K_0(4,3)g[x(t_3),t_3]\psi(3)\,dx_3\,dx_4 \quad (7\text{-}19)$$

This expression can be substituted into Eq. (7-18) to obtain the final result for the first Born approximation, $(i/\hbar)\langle\chi|\sigma|\psi\rangle_{S_0}$.

Transition elements will come up more frequently in succeeding chapters. In each example they can be evaluated in the straightforward manner which we have illustrated here. For that reason, very little of the material in the remainder of this chapter is really essential to the work that follows. Nevertheless, there are two reasons for the inclusion of this material in this book. First, it is possible to obtain a very general relation between transition elements. This relation might well serve as an alternative starting point for the definition of quantum mechanics. Second, for many people already familiar with the more conventional operator notation of quantum mechanics, it is helpful to have examples of the translation from the more customary representation into that which is used in this book, such as expressions of the form of Eq. (7-3).

With the rules for translation available, the subject matter of the later chapters, developed as it is from the path integral approach, can be appreciated in terms of more familiar symbolic concepts.

The relations discussed in the remainder of this chapter are independent of the form of the wave functions which describe either the initial or final state of the system, and they are used in defining the integral of the transition element. For this reason we shall abbreviate our notation by omitting any specific reference to these wave functions. Thus a transition element will be written as $\langle F\rangle_S$ instead of $\langle\chi|F|\psi\rangle_S$.

7-2 FUNCTIONAL DERIVATIVES

We are embarking on a mathematical development which leads to an interesting relation between transition elements. This relation finds its most elegant expression in terms of a mathematical idea, the functional derivative. Since this idea may not be familiar, we describe it in this section.

The functional $F[x(t)]$ gives a number for each function $x(t)$ that we may choose. We may ask: How much does this number change if we make a very small change in the argument function $x(t)$? Thus, for small $\eta(t)$, how much is $F[x(t) + \eta(t)] - F[x(t)]$? The effect to first order in η (assuming it exists, etc.) is some linear expression in η, say, $\int K(s)\eta(s)\,ds$. Then $K(s)$ is called the functional derivative of F with respect to variation of the function $x(t)$ at s. It is written $\delta F/\delta x(s)$. That is, to first order

$$F[x + \eta] = F[x] + \int \frac{\delta F}{\delta x(s)}\,\eta(s)\,ds + \cdots \tag{7-20}$$

This $\delta F/\delta x(s)$ depends on the function $x(t)$, of course, and also on the value of s. Thus it is a functional of $x(t)$ and a function of time s.

We may look at it another way. Suppose time is divided into very many steps of small interval ϵ, the values of the time being t_i $(t_{i+1} = \epsilon + t_i)$. The function $x(t)$ can now be specified approximately by giving the value x_i that it takes on at each of the times t_i. The functional $F[x(t)]$ now is a number depending on all the x_i; that is, it becomes an ordinary function of the variables x_i,

$$F[x(t)] \to F(\ldots, x_i, x_{i+1}, \ldots) \tag{7-21}$$

Now we can consider its derivative with respect to one of these variables x_i, $\partial F/\partial x_i$. Our functional derivative is just this partial, divided by ϵ and taken at the point $t_i = s$. That is,

$$\frac{\delta F}{\delta x(s)} \to \frac{1}{\epsilon}\frac{\partial F}{\partial x_i} \tag{7-22}$$

This we can see as follows. If we alter the path from $x(t)$ to $x(t) + \eta(t)$, we change all the x_i from x_i to $x_i + \eta_i$ [where $\eta_i = \eta(t_i)$], so that the first-order change in our function is

$$F(\ldots x_i + \eta_i, x_{i+1} + \eta_{i+1} \ldots) - F(\ldots, x_i, x_{i+1}, \ldots) = \sum_i \frac{\partial F}{\partial x_i}\eta_i \tag{7-23}$$

from the ordinary rules of partial differentiation. If now we call $(1/\epsilon)(\partial F/\partial x_i) = K_i$, the last sum is $\sum_i K_i\eta_i\epsilon$, which in the limit becomes $\int K(t)\eta(t)\,dt$. So if this limit exists as $\epsilon \to 0$, then it is equal to $\delta F/\delta x(s)$.

One can also use the ideas of differentials. Just as we can write

$$df = \sum_i \frac{\partial f}{\partial x_i}\,dx_i$$

so we can write for the first variation of any functional

$$\delta F = \int \frac{\delta F}{\delta x(s)}\,\delta x(s)\,ds \tag{7-24}$$

where $\delta x(s)$ is the differential change in path at $x(s)$.

Problem 7-1 If $S = \int_{t_1}^{t_2} L(\dot{x},x,t)\,dt$, show, for any s inside the range t_1 to t_2,

$$\frac{\delta S}{\delta x(s)} = -\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}}\right) + \frac{\partial L}{\partial x} \tag{7-25}$$

where the partial derivatives are evaluated at $t = s$.

Problem 7-2 If $F[x] = x(\tau)$, show that

$$\frac{\delta F}{\delta x(s)} = \delta(\tau - s) \tag{7-26}$$

Problem 7-3 Show that $\delta F/\delta j(\mathbf{r},s)$, where

$$F = \exp\left[\tfrac{1}{2}\int \cdots \int j(\mathbf{r}_1,t_1)j(\mathbf{r}_2,t_2)R(\mathbf{r}_1 - \mathbf{r}_2, t_1 - t_2)\, d^3\mathbf{r}_1\, d^3\mathbf{r}_2\, dt_1\, dt_2\right]$$

is given by

$$\frac{\delta F}{\delta j(\mathbf{r},s)} = \left[- \int R(\mathbf{r} - \mathbf{r}', s - t')j(\mathbf{r}',t')\, d\mathbf{r}'\, dt' \right] F \tag{7-27}$$

Note that the function $j(\mathbf{r},t)$ is a function of the four variables (x,y,z,t). Thus the single coordinate s, as used in Eq. (7-14), for example, must be replaced by the set of coordinates (x,y,z,t) in specifying the point at which the functional derivative is evaluated.

The general relation between functionals which we mentioned at the beginning of the preceding section may be obtained by trying to develop a formula for the transition element of $\delta F/\delta x(s)$. This we can do most easily in this way. Consider

$$\langle F\rangle_S = \int F[x(t)]e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) \tag{7-28}$$

Now in the integral over paths substitute $x(t) + \eta(t)$ for the variable $x(t)$. For fixed $\eta(t)$, $\mathfrak{D}[x(t) + \eta(t)] = \mathfrak{D}x(t)$ [since $d(x_i + \eta_i) = dx_i$]. But the integral is unchanged by a substitution of its variable. Hence

$$\begin{aligned}\langle F\rangle_S &= \int F[x(t) + \eta(t)]e^{(i/\hbar)S[x(t)+\eta(t)]}\, \mathfrak{D}x(t)\\ &= \int F[x(t)]e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) + \int \left[\int \frac{\delta F}{\delta x(s)}\eta(s)\, ds\right] e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t)\\ &\quad + \frac{i}{\hbar}\int F[x(t)]\left[\int \frac{\delta S}{\delta x(s)}\eta(s)\, ds\right] e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) + \cdots\end{aligned} \tag{7-29}$$

expanding the exponential and displaying only to first order. The zero-order term is exactly $\langle F\rangle_S$ again, so the remaining terms must all vanish. In particular, the first-order term must vanish for any $\eta(s)$, so that we conclude the relation

$$\left\langle\frac{\delta F}{\delta x(s)}\right\rangle_S = -\frac{i}{\hbar}\left\langle F\,\frac{\delta S}{\delta x(s)}\right\rangle_S \tag{7-30}$$

This general relation has many important consequences.

It would be possible to use Eq. (7-30) as a starting point to define the laws of quantum mechanics. One could work backward to reproduce, for example, Eq. (7-6). If some generalization of quantum mechanics is desired, one might suppose such a generalization is included in the action S appearing in the term $e^{iS/\hbar}$, or perhaps start with a form like Eq. (7-30) and introduce modifications with the help of the differential notation. Julian Schwinger has been investigating the formulation of quantum mechanics suggested by Eq. (7-30).

We can see how the relation of Eq. (7-30) comes about in another way by imagining our time split into intervals ϵ and functionals replaced by functions of the points x_i corresponding to t_i. Then consider the path integral

$$\int \frac{\partial F}{\partial x_k} e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) \tag{7-31}$$

where t_k is some intermediate time not at either end point. The path integral is simply an integral over all the points x_i. So we integrate by parts to get

$$\int \frac{\partial F}{\partial x_k} e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) = -\frac{i}{\hbar}\int F \frac{\partial S}{\partial x_k} e^{(i/\hbar)S[x(t)]}\, \mathfrak{D}x(t) \tag{7-32}$$

dropping the integrated part.

Problem 7-4 Discuss why the integrated part vanishes.

The result is

$$\left\langle \frac{\partial F}{\partial x_k} \right\rangle_S = -\frac{i}{\hbar}\left\langle F \frac{\partial S}{\partial x_k} \right\rangle_S \tag{7-33}$$

which is the same content as Eq. (7-30).

It is better to write these relations as differentials,

$$\langle \delta F \rangle_S = -\frac{i}{\hbar}\langle F\, \delta S \rangle_S \tag{7-34}$$

for then the specific variables on which F and S depend need not be indicated.

Problem 7-5 Argue that Eq. (7-34) may be misleading, for Eq. (7-33) applies only to rectangular coordinates. Do this by studying the corresponding relation where spherical coordinates, for example, are used and we wish to find $\langle \partial F/\partial r_k \rangle_S$.

7-3 TRANSITION ELEMENTS OF SOME SPECIAL FUNCTIONALS

The relation of Eq. (7-34) has many interesting implications. In this section we shall investigate some of them. We shall take the special case of a one-dimensional particle moving in a potential $V[x(t)]$.

Suppose the action over the path of the particle is given by

$$S = \int_{t_1}^{t_2} \left\{ \frac{m\dot{x}^2}{2} - V[x(t)] \right\} dt \tag{7-35}$$

Upon application of the small variation $\delta x(t)$ to each path there results (to first order)

$$\delta S = -\int_{t_1}^{t_2} [m\ddot{x} + V'(x)]\,\delta x(t)\,dt \tag{7-36}$$

Using Eq. (7-34), we have

$$\langle \delta F \rangle_S = -\frac{i}{\hbar}\left\langle F \int_{t_1}^{t_2} [m\ddot{x} + V'(x)]\delta x(t)\,dt \right\rangle \tag{7-37}$$

Alternatively, we could return to the point of view used in developing Eq. (7-33). That is, we imagine time divided into small slices of length ϵ. In this case the action S can be written as

$$S = \sum_{i=1}^{N-1} \left[m\frac{(x_{i+1} - x_i)^2}{2\epsilon} - V(x_i)\epsilon \right] \tag{7-38}$$

If we select a particular time t_k and, as before, let x_k be the associated position of a path, then

$$\frac{\partial S}{\partial x_k} = m\left(\frac{x_{k+1} - x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon} \right) + V'(x_k)\epsilon \tag{7-39}$$

Upon application of Eq. (7-33) there results

$$\left\langle \frac{\partial F}{\partial x_k} \right\rangle_S = -\frac{i\epsilon}{\hbar}\left\langle F\left[m\left(\frac{x_{k+1} - 2x_k + x_{k-1}}{\epsilon^2} \right) + V'(x_k) \right] \right\rangle \tag{7-40}$$

In this last expression the factor involving an ϵ^2 in the denominator is actually the acceleration $\ddot{x}$ evaluated at the time t_k. Thus Eq. (7-40) is just a special example of Eq. (7-37). In particular, it corresponds to Eq. (7-37) if $\delta x(t)$ is zero for all $t \neq t_k$. If $\delta x(t)$ is assigned the value $\epsilon \cdot \delta x_k \cdot \delta(t - t_k)$, then Eq. (7-40) results. Actually Eq. (7-40), since it is true for all k, is completely equivalent to Eq. (7-37) in a more detailed notation.

In Eq. (7-37) suppose we choose the special function $F = 1$. Then $\delta F = 0$ and we have

$$-\frac{i}{\hbar}\left\langle \int [m\ddot{x} + V'(x)]\,\delta x(t)\,dt \right\rangle = 0 \tag{7-41}$$

Since this result must hold for any arbitrary choice of $\delta x(t)$, it must be that

$$\langle m\ddot{x}\rangle = -\langle V'(x)\rangle \tag{7-42}$$

at all values of time. This is the quantum-mechanical analogue of Newton's law. Making use of the classical analogue for a transition element, described in Sec. 7-1, this result says that the weighted "average" of the mass times acceleration at any time "averaged" over all paths with the weight $e^{iS/\hbar}$ is equal to the weighted "average" of the force (negative gradient of the potential) at the same time.

As another example, suppose F is some arbitrary nonzero functional of all position variables *except* x_k. Then the left-hand side of Eq. (7-40) is zero and there results (since $\partial F/\partial x_k = 0$)

$$\left\langle F(x_1,x_2,\ .\ .\ .\ ,x_{k-1},x_{k+1},\ .\ .\ .\ ,x_N)\left[m\,\frac{x_{k+1} - 2x_k + x_{k-1}}{\epsilon^2} + V'(x_k)\right]\right\rangle = 0 \tag{7-43}$$

This equation says that the transition element of $m\ddot{x} + V'(x)$, averaged over all paths, is zero at t_k even if these paths are weighted with an arbitrary functional, so long as the functional is independent of the position of the path at the time t_k of interest.

Suppose, however, the functional does depend upon the position of the path at the time of interest. In particular, suppose simply that the functional F is x_k. Applying Eq. (7-40), we have

$$\begin{aligned}\langle 1\rangle &= \frac{i\epsilon}{\hbar}\left\langle mx_k\,\frac{x_{k+1} - 2x_k - x_{k-1}}{\epsilon^2} + x_kV'(x_k)\right\rangle \\ &= \frac{i}{\hbar}\left\langle mx_k\left(\frac{x_{k+1} - x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon}\right) + \epsilon x_kV'(x_k)\right\rangle\end{aligned} \tag{7-44}$$

If we suppose that the potential V is a smooth function, then in the limit as $\epsilon \to 0$ we find that $\epsilon x_kV'(x_k)$ becomes negligible in comparison with the remaining terms. The result is

$$\left\langle m\,\frac{x_{k+1} - x_k}{\epsilon}\,x_k\right\rangle - \left\langle x_km\,\frac{x_k - x_{k-1}}{\epsilon}\right\rangle = \frac{\hbar}{i}\langle 1\rangle \tag{7-45}$$

This last equation involves the product of position variables x and momentum variables $m\dot{x}$. In the first term the momentum is evaluated first as a linear average corresponding to the time $t_k + \epsilon/2$, and the position is taken at t_k. In the second term the position is again taken at t_k, but the momentum corresponds to the time $t_k - \epsilon/2$. Thus this equation says that the transition element of a product of position and momentum depends upon the order in time of these two quantities.

Later on, when we make a translation into the more usual operator notion, we shall see that both the operator equation of motion, corresponding to Eq. (7-42), and the operator commutation laws of Eq. (7-45) have been derived from the same fundamental relation, Eq. (7-34).

We can derive a further result from Eq. (7-45) which will give us a better idea of the characteristics of the paths which are important in quantum mechanics. Consider the two terms

$$\left\langle x_k m \frac{x_k - x_{k-1}}{\epsilon} \right\rangle \tag{7-46}$$

and

$$\left\langle x_{k+1} m \frac{x_{k+1} - x_k}{\epsilon} \right\rangle \tag{7-47}$$

These two terms differ from each other only in order ϵ, since they are the same quantity calculated at two times differing by the interval ϵ. Thus we are justified in substituting Eq. (7-47) for the second term in Eq. (7-45). The result is

$$\left\langle m \frac{x_{k+1} - x_k}{\epsilon} (x_k - x_{k+1}) \right\rangle = \frac{\hbar}{i} \langle 1 \rangle \tag{7-48}$$

Alternatively, we can write this as

$$\left\langle \left(\frac{x_{k+1} - x_k}{\epsilon} \right)^2 \right\rangle = -\frac{\hbar}{im\epsilon} \langle 1 \rangle \tag{7-49}$$

This equation says that the transition element of the square of the velocity is of the order $1/\epsilon$, and thus becomes infinite as ϵ approaches zero. This result implies that the important paths for a quantum-mechanical particle are not those which have a definite slope (or velocity) everywhere, but are instead quite irregular on a very fine scale, as indicated by the sketch of Fig. 7-1. In fact, these irregularities are such that the "average" square velocity does not exist, where we have used the classical analogue in referring to an "average."

If some average velocity is defined for a short time interval Δt, as, for example, $[x(t + \Delta t) - x(t)]/\Delta t$, the "mean" square value of this is $-\hbar/(im\,\Delta t)$. That is, the "mean" square value of a velocity averaged over a short time interval is finite, but its value becomes larger as the interval becomes shorter.

It appears that quantum-mechanical paths are very irregular. However, these irregularities average out over a reasonable length of time to produce a reasonable drift, or "average" velocity, although for short intervals of time the "average" value of the velocity is very high.

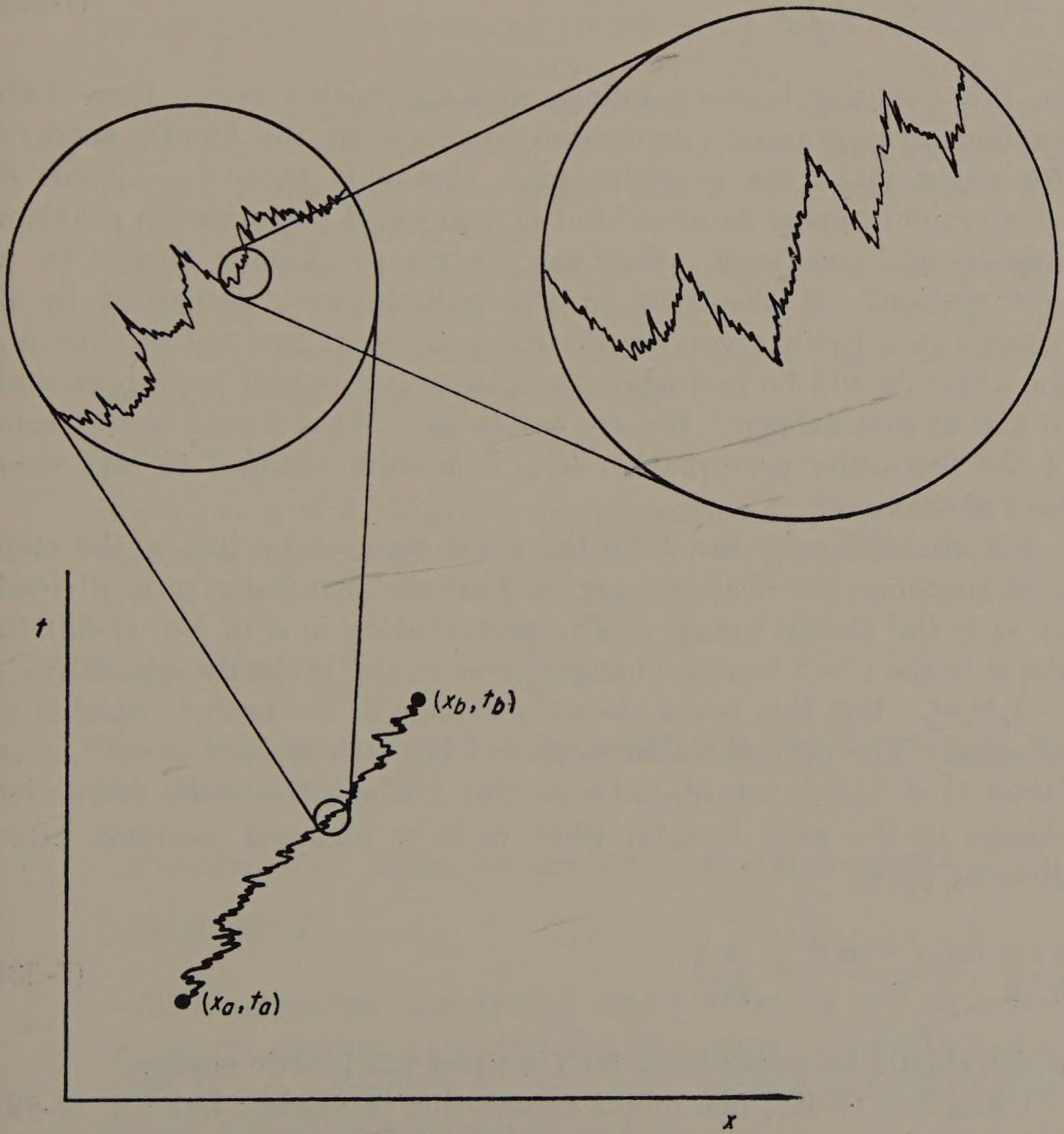


Fig. 7-1 Typical paths of a quantum-mechanical particle are highly irregular on a fine scale, as shown in the sketch. Thus, although a mean velocity can be defined, no mean-square velocity exists at any point. In other words, the paths are nondifferentiable.

Problem 7-6 Show, for a particle moving in three-dimensional space x, y, z,

$$\langle (x_{k+1} - x_k)^2 \rangle = \langle (y_{k+1} - y_k)^2 \rangle = \langle (z_{k+1} - z_k)^2 \rangle = -\frac{\hbar\epsilon}{im} \tag{7-50}$$

$$\begin{aligned}\langle (x_{k+1} - x_k)(y_{k+1} - y_k) \rangle &= \langle (x_{k+1} - x_k)(z_{k+1} - z_k) \rangle \\ &= \langle (y_{k+1} - y_k)(z_{k+1} - z_k) \rangle = 0\end{aligned} \tag{7-51}$$

It will not do to write the transition element of the kinetic energy simply as

$$\frac{1}{2}\left\langle m\left(\frac{x_{k+1} - x_k}{\epsilon}\right)^2\right\rangle \tag{7-52}$$

for this quantity becomes infinite as ϵ approaches zero. How shall we find an appropriate expression to represent the kinetic energy? We might make the heuristic guess that only those functionals F which might appear in some kind of a physical perturbation problem may be of importance. How can we get the kinetic energy by a perturbation? If the mass of the particle were perturbed by a factor $1 + \eta$ (with η very small) for some short interval of time Δt, the action would be perturbed by $\eta\,\Delta t(m/2)\dot{x}^2$, which is proportional to the kinetic energy. We are led to ask: What would be the form of the first-order perturbation $\langle\sigma\rangle_{S_0}$ if m were changed to $m(1 + \eta)$ for a short time?

For simplicity we can take the short time to be just ϵ, the step used to define the time spacing, so that the first-order term divided by $\epsilon\eta$ is the kinetic energy. The perturbation in S of Eq. (7-38) (if the m in the $i = k$ term is changed to $m + \eta m$) is clearly $\epsilon\eta(m/2)(x_{k+1} - x_k)^2/\epsilon^2$. But this is not the only change in the path integral if m changes. The normalization factors A for each m vary as $m^{+1/2}$, so a factor $(1 + \frac{1}{2}\eta)$ is introduced from this. Hence the entire first-order change in the path integral when m is so changed becomes, after dividing by $\eta\epsilon$,

$$\frac{i}{\hbar}\left(\frac{m}{2}\frac{(x_{k+1} - x_k)^2}{\epsilon^2} + \frac{\hbar}{2i\epsilon}\right) \tag{7-53}$$

which should be satisfactory for $i/\hbar$ times the kinetic energy.

Using Eq. (7-49), one might expect this to vanish; but Eq. (7-49) is valid only as $\epsilon \to 0$ to the order $1/\epsilon$. The quantity in Eq. (7-53) is, in fact, finite as $\epsilon \to 0$. The expression can be rewritten by expanding the quadratic term. In Eq. (7-40) let F be $x_{k+1} - x_k$. If terms of lowest order in ϵ are kept, the result is

$$\left\langle \frac{m}{2}\left(\frac{x_{k+1}-x_k}{\epsilon}\right)\left(\frac{x_k-x_{k-1}}{\epsilon}\right)\right\rangle = \left\langle \frac{m}{2}\left(\frac{x_{k+1}-x_k}{\epsilon}\right)^2\right\rangle + \frac{\hbar}{2i\epsilon}\langle 1\rangle \tag{7-54}$$

Thus we can define the left-hand side of Eq. (7-54) as the transition element of the kinetic energy.

We see from this result that the easiest way to produce satisfactory transition elements involving powers of the velocities is to replace these powers by a product of velocities, each factor of which is taken at a slightly different time.

In simple problems the transition elements can sometimes be evaluated directly. For such problems the same results can also be obtained by using the relations among transition elements which we derived in Sec. 7-2. These relations may supply us with soluble differential equations for the transition elements. We shall give a few illustrations, but it will be readily seen that the examples for which the method works must be so simple that a direct evaluation would not really be much more difficult.

For our first example, consider the case of a free particle going from x_1 to x_2 in the total time interval T. Let us find the transition element of the position at the time t, that is, $x(t)$. Of course, this is some function of t and it is clear that

$$\langle x(0)\rangle = x_1\langle 1\rangle \qquad \langle x(T)\rangle = x_2\langle 1\rangle \tag{7-55}$$

Since any potentials acting on the particle are constant in space (i.e., no forces act), the second derivative of the transition element of position is zero in accordance with Eq. (7-42). Thus an integration gives

$$\langle x(t)\rangle = \left[x_1 + \frac{t}{T}(x_2 - x_1)\right]\langle 1\rangle \tag{7-56}$$

Note that the expression in the brackets is just the value of $x(t)$ along the classical path $\bar{x}(t)$.

Problem 7-7 Show for any action that is quadratic

$$\langle x(t)\rangle = \bar{x}(t)\langle 1\rangle \tag{7-57}$$

As a somewhat less trivial example, let us try to evaluate the transition element $\langle x(t)x(s)\rangle$ for the same free-particle conditions. Since this is a function of two times, we can write it as $f(t,s)$. The second derivative with respect to t is

$$\frac{\partial^2 f(t,s)}{\partial t^2} = \langle \ddot{x}(t)x(s)\rangle \tag{7-58}$$

This transition element can be worked out by substituting $F = x(s)$ into Eq. (7-40). For $s \neq t$, following the arguments leading to Eq. (7-42), the result is $-(1/m)\langle V'[x(t)]x(s)\rangle$; while for $s = t$, following the arguments leading to Eq. (7-44), we find that the transition element of Eq. (7-58) is of order $1/\epsilon$. In the limit of small ϵ we have

$$m\frac{\partial^2 f}{\partial t^2} = \langle m\ddot{x}(t)x(s)\rangle = \frac{\hbar}{i}\,\delta(t-s) - \langle V'[x(t)]x(s)\rangle \tag{7-59}$$

Since for our free particle the potential is independent of position, the second term on the right of Eq. (7-59) vanishes. The resulting equation may be solved by dividing the region of interest into two parts. For $t < s$

$$f = a(s)t + b(s) \tag{7-60}$$

while for $t > s$

$$f = A(s)t + B(s) \tag{7-61}$$

Thus the first derivative of the function f with respect to t jumps by the quantity $A(s) - a(s)$ as t goes from just below to just above s, and in accordance with Eq. (7-59), $A(s) - a(s) = \hbar/mi$.

The boundary conditions state that

$$\begin{aligned}\langle x(0)x(s)\rangle &= x_1\langle x(s)\rangle = x_1\bar{x}(s)\langle 1\rangle\\ \langle x(T)x(s)\rangle &= x_2\bar{x}(s)\langle 1\rangle\end{aligned} \tag{7-62}$$

This is not enough information to determine all of the four functions a, A, b, and B, but we can either make use of the relation

$$\frac{\partial^2 f}{\partial s^2} = \left(\frac{\hbar}{mi}\right)\delta(t-s) \tag{7-63}$$

obtained by differentiating f with respect to s, or else notice that $f(t,s)$ must be symmetric in t and s. One can conclude that the functions a, A, b, and B must all be linear in s. The boundary conditions are now sufficient to determine the solution. The result is

$$\begin{aligned}\langle x(t)x(s)\rangle &= \left[\bar{x}(t)\bar{x}(s) + \frac{\hbar}{miT}\,s(T-t)\right]\langle 1\rangle \qquad \text{for } s < t\\ &= \left[\bar{x}(t)\bar{x}(s) + \frac{\hbar}{miT}\,t(T-s)\right]\langle 1\rangle \qquad \text{for } t < s\end{aligned} \tag{7-64}$$

That this result is right can be seen by inspection. The product of two classical paths taken at different times, $\bar{x}(t)\bar{x}(s)$, is the solution of the homogeneous equations obtained by setting the right-hand sides of Eqs. (7-59) and (7-63) equal to zero, which satisfies the necessary boundary conditions. The last terms on the right of the pair of equations (7-64) are the special solutions of the inhomogeneous equations (7-59) and (7-63), which are zero at the end points.

The transition element of the product of two positions taken at two different times contains more than just the product of the two corresponding positions along the classical path. There is a small additional term which is purely quantum-mechanical in nature. This additional term is consistent with our picture of quantum-mechanical motion. Even though the particle moving between fixed end points will be found on the average along the classical path, it has an amplitude for motion along all alternative paths. This fact must be remembered when considering the transition element of the product of positions at two different times. All the possible positions among all the various alternatives must be accounted for in the transition element, and this accounting introduces the extra term. Only at the specified end points are no other alternatives possible.

We can better understand the significance of this result if we make use once again of the terminology from our classical analogue. Suppose the path of the particle goes through a particularly large value of x at some time s. Then the "average" value of x at a later time t is not just the ordinary average $\bar{x}(t)$. There is a correlation with the previous large deflection. Therefore, the "average" product is not just the product of "averages."

In this and other applications of the classical analogue, we remember that the "average" referred to is defined with the help of the weighting function $e^{iS/\hbar}$. This weighting function is not positive definite, and is in fact complex. Thus we develop such purely quantum-mechanical results as that of Eqs. (7-64) wherein the extra correlation term is pure imaginary!

Problem 7-8 Find the transition element $x(t)x(s) = f(t,s)$ when the potential is not constant but, rather, corresponds to that of a forced harmonic oscillator. Do this by obtaining differential equations for f and trying the solution

$$\langle f(t,s)\rangle = [\bar{x}(t)\bar{x}(s) + g(t,s)]\langle 1\rangle \tag{7-65}$$

Obtain an equation for $g(t,s)$ showing that g is independent of the end-point values x_1 and x_2 and of the forcing function [derivative of the potential) $\gamma(t)$]. Show in general that, with $T = t_2 - t_1$,

$$
\begin{aligned}
g(t,s) &= \frac{\hbar}{mi\omega \sin \omega T} \sin \omega s \sin \omega(T - t) \qquad s < t \\
&= \frac{\hbar}{mi\omega \sin \omega T} \sin \omega t \sin \omega(T - s) \qquad t < s
\end{aligned}
\tag{7-66}
$$

7-4 GENERAL RESULTS FOR QUADRATIC ACTIONS

Evidently if the action S is a quadratic form, transition elements of many functionals can be determined readily. This suggests that we extend our consideration into a somewhat more general class of functionals. The technique to be used is the same as that described in Sec. 3-5. For example, we note that with a quadratic action S we can easily evaluate the transition element of $\exp (i/\hbar)\int f(t)x(t)\,dt$, where $f(t)$ is any arbitrary function of time. The transition element of such a functional can be written as

$$
\left\langle e^{(i/\hbar)\int f(t)x(t)\,dt} \right\rangle = \int_a^b e^{(i/\hbar)[S+\int f(t)x(t)\,dt]}\,\mathcal{D}x(t)
\tag{7-67}
$$

If the original action S is gaussian, then so is the action

$S' = S + \int f(t)x(t)\,dt$

Thus the path integrals on the right of Eq. (7-67) can be carried out by the methods of Sec. 3-5. If S'_{cl} is the extremum of the action S', then the factor $\exp (iS'_{cl}/\hbar)$ can be extracted as a factor for the path integral of Eq. (7-67). The remaining factor is a path integral over the paths $y(t)$, which run from zero to zero during the allowed time interval. (We set $x = \bar{x} + y$, where $\bar{x}$ is the classical path corresponding to the extremum of the action.)

The integral over the paths y does not depend upon the function $f(t)$, since this function appears in the action S' multiplying only a linear term in $x(t)$, and we have seen [cf. Eq. (3-49)] that the remaining path integral involves only the quadratic parts of S' which are not more than the quadratic parts of S. This means that the path integral on the right-hand side of Eq. (7-67) can be reduced to an exponential function multiplied by the transition element $\langle 1 \rangle$. The result is

$$
\left\langle \exp\left[\frac{i}{\hbar}\int f(t)x(t)\,dt\right]\right\rangle = \left\{\exp\left[\frac{i}{\hbar}(S'_{cl} - S_{cl})\right]\right\}\langle 1 \rangle
\tag{7-68}
$$

Once the extremum S'_{cl} has been evaluated, the extremum S_{cl} can be obtained from it by setting $f(t)$ identically equal to zero. The action of the forced harmonic oscillator, defined in Eq. (3-68), is a special case of this action S'_{cl}.

Problem 7-9 Use this result to show that if S corresponds to an harmonic oscillator

$$S = \frac{m}{2}\int \dot{x}^2\,dt - \frac{m\omega^2}{2}\int x^2\,dt$$

then

$$\left\langle \exp\left[\frac{i}{\hbar}\int f(t)x(t)\,dt\right]\right\rangle = \langle 1\rangle \left(\exp\left\{\left[\frac{i}{\hbar}\frac{m\omega}{2\sin\omega(t_2-t_1)}\right]\right.\right.$$
$$\times\left[\frac{2x_2}{m\omega}\int_{t_1}^{t_2} f(t)\sin\omega(t-t_1)\,dt + \frac{2x_1}{m\omega}\int_{t_1}^{t_2} f(t)\sin\omega(t_2-t)\,dt\right.$$
$$\left.\left.\left. - \frac{2}{m^2\omega^2}\int_{t_1}^{t_2}\int_{t_1}^{t} f(t)f(s)\sin\omega(t_2-t)\sin\omega(t-t_1)\,ds\,dt\right]\right\}\right)$$

where x_1, x_2 are the initial and final coordinates of the oscillator.

From the transition element given by Eq. (7-68) we can obtain the transition element of $x(t)$ itself by another method. Suppose we differentiate Eq. (7-68) with respect to $f(t)$. The result is

$$\left\langle x(t)\exp\left[\frac{i}{\hbar}\int fx\,dt\right]\right\rangle = \frac{\hbar}{i}\frac{\delta}{\delta f(t)}\left\{\exp\left[\frac{i}{\hbar}(S'_{cl}-S_{cl})\right]\right\}\langle 1\rangle$$
$$= \frac{\delta S'_{cl}}{\delta f(t)}\left\{\exp\left[\frac{i}{\hbar}(S'_{cl}-S_{cl})\right]\right\}\langle 1\rangle \tag{7-69}$$

Therefore, by evaluating both sides when $f(t) \equiv 0$, we obtain

$$\langle x(t)\rangle = \langle 1\rangle \left.\frac{\delta S'_{cl}}{\delta f(t)}\right|_{f=0} \tag{7-70}$$

We can continue this process to get the second derivative as

$$\langle x(t)x(s)\rangle = \left(\frac{\hbar}{i}\right)^2 \frac{\delta^2}{\delta f(t)\,\delta f(s)}\exp\left[\frac{i}{\hbar}(S'_{cl}-S_{cl})\right]\Big|_{f=0}\langle 1\rangle$$
$$= \langle 1\rangle\left[\frac{\hbar}{i}\frac{\delta^2 S'_{cl}}{\delta f(t)\,\delta f(s)} + \frac{\delta S'_{cl}}{\delta f(t)}\frac{\delta S'_{cl}}{\delta f(s)}\right]\Big|_{f=0} \tag{7-71}$$

Actually, since S'_{cl} is quadratic only in f [cf. Eq. (3-66)], the transition element of a factor of any number of x's can be directly evaluated in terms of $\delta S'_{cl}/\delta f(t)$ and the quantity $\delta^2 S_{cl}/\delta f(t)\,\delta f(s)$, which is independent of f. This explains the form of Eqs. (7-64) and (7-65) and permits the transition element of a factor of three x's to be written down.

Problem 7-10 Show, for any quadratic functional, if we write $\langle x(t)\rangle = \bar{x}(t)\langle 1\rangle$ and $\langle x(t)x(s)\rangle = [(\bar{x}(t)\bar{x}(s) + g(t,s)]\langle 1\rangle$, that

$$\langle x(t)x(s)x(u)\rangle = [(\bar{x}(t)\bar{x}(s)\bar{x}(u) + \bar{x}(t)g(s,u) + \bar{x}(s)g(t,u) + \bar{x}(u)g(t,s)]\langle 1\rangle$$

Find the transition element of the product of four x's. [*Suggestion:* Since $S'_{cl} - S_{cl}$ is quadratic in f and zero for $f = 0$, it must have the mathematical form $S'_{cl} - S_{cl} = \frac{1}{2}\int\int f(t)f(s)g(t,s)\,dt\,ds + \int \bar{x}(t)f(t)\,dt$, where g and $\bar{x}$ are some functions.]

7-5 TRANSITION ELEMENTS AND THE OPERATOR NOTATION

In this and the following sections we shall see how transition elements look in the conventional notation of wave functions and operators. This will help the reader who is familiar with that form of expression to relate the results of path integral calculations to other results that he already knows.

If F is a function only of x at a single time, say, the function $V(x_k)$ at time t_k, we know from Eq. (7-10) how to evaluate its transition element. Similarly, if F depends on the value of $x(t)$ at two different times, Eq. (7-15) tells us what to do.

Let us consider next the case that F represents the momentum at time t_k and make use of the approximation that the time axis is cut up into slices of length ϵ. Thus

$$F = \frac{m}{\epsilon}(x_{k+1} - x_k) \tag{7-72}$$

Then we have

$$\left\langle \chi \left| \frac{m}{\epsilon}(x_{k+1} - x_k) \right| \psi \right\rangle_S = \frac{m}{\epsilon}(\langle \chi | x_{k+1} | \psi \rangle_S - \langle \chi | x_k | \psi \rangle_S) \tag{7-73}$$

The right-hand side of Eq. (7-73) can be written as

$$\frac{m}{\epsilon}\left[\int \chi^*(x, t+\epsilon)x\psi(x, t+\epsilon)\,dx - \int \chi^*(x,t)x\psi(x,t)\,dx\right] \tag{7-74}$$

Now making use of the wave equation

$$\psi(x, t+\epsilon) = \psi(x,t) + \epsilon\frac{\partial\psi}{\partial t} = \psi - \frac{i\epsilon}{\hbar}H\psi \tag{7-75}$$

$$\chi^*(x, t+\epsilon) = \chi^*(x,t) + \epsilon\frac{\partial\chi^*}{\partial t} = \chi^* + \frac{i\epsilon}{\hbar}(H\chi)^* \tag{7-76}$$

from Prob. 4-3, if H is the hamiltonian belonging to the S. Therefore, to first order in ϵ

$$\int \chi^*(x, t+\epsilon)x\psi(x, t+\epsilon)\,dx = \int \chi^*(x,t)x\psi(x,t)\,dx \\ - \frac{i\epsilon}{\hbar}\left\{\int \chi^*(x,t)x[H\psi(x,t)]\,dx - \int [H^*\chi^*(x,t)]x\psi(x,t)\,dx\right\} \tag{7-77}$$

By Eq. (4-30) this last integral can be written as $\int\chi^*(x,t)[Hx\psi(x,t)]\,dx$, or more simply we have

$$\langle\chi|m\dot{x}|\psi\rangle = -\frac{im}{\hbar}\int \chi^*(xH - Hx)\psi\,dx \tag{7-78}$$

using the operator notation. This is the same as

$$-\frac{i}{\hbar}m\int \chi^*\frac{\hbar^2}{m}\frac{\partial\psi}{\partial x}\,dx = \int \chi^*\frac{\hbar}{i}\frac{\partial\psi}{\partial x}\,dx \tag{7-79}$$

where we have used the result of Prob. 4-4. The operator $(\hbar/i)\,\partial/\partial x$ is called the momentum operator or, more specifically, the operator representing momentum in the x direction. We already see why. Constructing the transition element of $m\dot{x}$ is equivalent to putting the operator $(\hbar/i)\,\partial/\partial x$ between χ^* and ψ, just as constructing the transition element of x is equivalent to putting x between χ^* and ψ. These relations can be understood, perhaps with greater clarity, if we go over to the momentum representation. If

$$\begin{aligned}\chi(p) &= \int_{-\infty}^{\infty} \chi(x)e^{-(i/\hbar)px}\,dx \\ \psi(p) &= \int_{-\infty}^{\infty} \psi(x)e^{-(i/\hbar)px}\,dx\end{aligned} \tag{7-80}$$

are the momentum representations of χ and ψ, one can show that

$$\int_{-\infty}^{\infty} \chi^*(x)\frac{\hbar}{i}\frac{\partial\psi(x)}{\partial x}\,dx = \int_{-\infty}^{\infty} \chi^*(p)p\psi(p)\,dp \tag{7-81}$$

Problem 7-11 Show this.

Another way to see this relation is the following. Consider the transition amplitude given by

$$\langle\chi|1|\psi\rangle = \int\int \chi^*(x_N,t_N)K(x_N,t_N;x_1,t_1)\psi(x_1,t_1)\,dx_1\,dx_N \tag{7-82}$$

Now suppose the whole x_1 axis is shifted to the right by a small amount Δ. Calling the new axis x_1', we have

$$x_1 = x_1' - \Delta \tag{7-83}$$

Using this new variable rather than the old x_1 will not alter the transition amplitude of Eq. (7-82). It becomes

$$\langle \chi|1|\psi\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{x_1'}^{x_N} \chi(x_N,t_N) \exp\left\{\left(\frac{i}{\hbar}\right) \sum_{k=2}^{N-1} S[x_{k+1},t_{k+1};x_k,t_k] + \left(\frac{i}{\hbar}\right) S[x_2,\, t_2;\, x_1' - \Delta,\, t]\right\} \psi(x_1' - \Delta,\, t)\, \mathfrak{D}x(t)\, dx_1'\, dx_2 \tag{7-84}$$

where the path integral for the kernel has been written out explicitly, using the methods of Eq. (2-22).

Next, we expand $S[x_2,\, t_2;\, x_1' - \Delta,\, t]$ and $\psi(x_1' - \Delta,\, t)$ in Taylor series and keep only the first-order terms. In particular, the exponential function becomes

$$\exp\left\{\sum_{k=2}^{N-1} \left(\frac{i}{\hbar}\right) S[x_{k+1},t_{k+1};x_k,t_k]\right\}\left\{1 - \left(\frac{i}{\hbar}\right) \Delta \frac{\partial}{\partial x_1'} S[x_2,t_2;x_1',t_1]\right\} \tag{7-85}$$

We may drop the prime notation in the integral defining the transition amplitude, since x_1' is a variable of integration. The form of Eq. (7-84) now becomes

$$\langle \chi|1|\psi\rangle = \iint \chi^*(2)K(2,1)\psi(1)\, dx_1\, dx_2 - \frac{i}{\hbar}\Delta \iint \chi^*(2)K(2,1) \times \left\{\psi_1 \frac{\partial}{\partial x_1} S[x_2,t_2;x_1,t_1] + \frac{\hbar}{i}\frac{\partial}{\partial x_1}\psi(x_1,t_1)\right\} dx_1\, dx_2 \tag{7-86}$$

where we retain the notation that the point x_2 is spaced along the path $x(t)$ only by the short time interval ϵ from the point x_1 and $t_2 = t_1 + \epsilon$.

The first term on the right of Eq. (7-86) is identical to the transition amplitude on the left. This means that the remaining term must be zero. But this remaining term is a combination of two transition elements. Thus

$$\left\langle \chi \left| -\frac{\partial}{\partial x_1} S[x_2,\, t_1 + \epsilon;\, x_1,\, t_1] \right| \psi \right\rangle = \left\langle \chi|1| \frac{\hbar}{i} \frac{\partial\psi(x_1,t_1)}{\partial x_1} \right\rangle \tag{7-87}$$

In the convention of Eq. (2-22) we use the classical action along each of the short elements of the path. Thus the action $S[2,1]$ appearing in Eq. (7-87) is the classical action for the initial path element. Its negative derivative with respect to x_1 is the classical definition of the momentum at x_1. So we can write

$$\langle \chi|p_1|\psi\rangle = \left\langle \chi|1| \frac{\hbar}{i} \frac{\partial\psi}{\partial x_1}\right\rangle \tag{7-88}$$

which is the same result as that obtained in Eqs. (7-78) and (7-79).

Sometimes working with a complicated S that results perhaps from the partial elimination of interacting parts, we would like to identify the functional $p(t)$ which corresponds to the momentum at time t. The work of the preceding paragraph suggests a general definition. The first-order change in the transition amplitude $\langle\chi|1|\psi\rangle$, if all coordinates corresponding to times previous to t are shifted by $-\Delta$, is this Δ times $\langle\chi|p(t)|\psi\rangle$. From this principle the momentum functional may be found for an arbitrarily complicated S. In a like manner, the hamiltonian or energy functional can be defined by shifting the time variables, as we shall describe in Sec. 7-7.

Problem 7-12 Show, if V is any function of position only, that

$$\left\langle\chi\left|\frac{dV}{dt}\right|\psi\right\rangle = \left\langle\chi\left|\frac{V(x_{k+1}) - V(x_k)}{\epsilon}\right|\psi\right\rangle$$

$$= \frac{i}{\hbar}\int_{-\infty}^{\infty} \chi^*(HV - VH)\psi\,dx \tag{7-89}$$

Consider the case that V is also a function of the time. Show that the transition element of dV/dt is equivalent to the transition element of the operator $(i/\hbar)(HV - VH) + \partial V/\partial t$.

Problem 7-13 Show that

$$\langle\chi|m\ddot{x}|\psi\rangle = \frac{i}{\hbar}\int_{-\infty}^{\infty} \chi^*(Hp - pH)\psi\,dx \tag{7-90}$$

and argue for any quantity A, given in terms of an operator or otherwise, that dA/dt is equivalent to $\partial A/\partial t + (i/\hbar)(HA - AH)$.

Next we consider an expression F involving two quantities evaluated in rapid succession, such as

$$F = \frac{m(x_{k+1} - x_k)}{\epsilon}\,x_k \tag{7-91}$$

This evidently gives

$$\langle\chi|F|\psi\rangle = \frac{1}{\epsilon}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi^*(x;t+\epsilon)mxK(x,t+\epsilon;y,t)y\psi(y,t)\,dy\,dx$$

$$- \frac{1}{\epsilon}\int \chi^*(x,t)mx^2\psi(x,t)\,dx \tag{7-92}$$

where $t = t_k$. In developing Eq. (4-12) from Eq. (4-2) we saw

$$\int_{-\infty}^{\infty} K(x,t+\epsilon;y,t)f(y)\,dy = f(x) + \frac{i\epsilon}{\hbar}Hf(x) \tag{7-93}$$

so that the first integral in Eq. (7-92) is

$$\frac{1}{\epsilon}\int_{-\infty}^{\infty} \chi^*(x, t+\epsilon)mx\left(1+\frac{i\epsilon}{\hbar}H\right)x\psi(x,t)\,dx \tag{7-94}$$

Expressing χ^* by Eq. (7-76) and using the hermitian property of H, we find that this integral is

$$\begin{aligned}\frac{1}{\epsilon}\int_{-\infty}^{\infty} &\chi^*(x,t)\left(1-\frac{i\epsilon H}{\hbar}\right)mx\left(1+\frac{i\epsilon}{\hbar}H\right)x\psi(x,t)\,dx\\ &=\frac{1}{\epsilon}\int_{-\infty}^{\infty}\chi^*(x,t)mx^2\psi(x,t)\,dx+\frac{i}{\hbar}\int_{-\infty}^{\infty}\chi^*(x,t)m(xH-Hx)x\psi(x,t)\,dx\end{aligned} \tag{7-95}$$

Thus finally

$$\begin{aligned}\left\langle\chi\left|m\frac{x_{k+1}-x_k}{\epsilon}x_k\right|\psi\right\rangle&=\frac{i}{\hbar}\int\chi^*(x,t)m(xH-Hx)x\psi(x,t)\,dx\\&=\int\chi^*(x,t)px\psi(x,t)\,dx\end{aligned} \tag{7-96}$$

the last step following from Eq. (7-78). This is an example of the general rule: In writing the integral definition of the transition element for a set of quantities corresponding to a succession of times, the corresponding operators are written in order from right to left, according to the order in time of the original transition element. If there is a finite time interval Δt between them, a K, or alternatively the operator $e^{-(i/\hbar)S\,\Delta t}$, must be inserted. (For an example, see Prob. 7-16.) As the time interval ϵ between two successive quantities approaches zero the K approaches a δ function and the rule results.

Problem 7-14 Show that the transition amplitude of $(m/\epsilon)(x_{k+1}-x_k)f(x_{k+1})$ is equivalent to that of $(f\cdot p)$.

Problem 7-15 Show that the rule works for two successive momenta, that is,

$$\begin{aligned}\left\langle\chi\left|m\frac{x_{k+1}-x_k}{\epsilon}m\frac{x_k-x_{k-1}}{\epsilon}\right|\psi\right\rangle&=\int\int_{-\infty}^{\infty}\chi^*(y,t)pp\psi(x,t)\,dx\,dy\\&=-\hbar^2\int_{-\infty}^{\infty}\chi^*(y,t)\frac{\partial^2}{\partial x^2}\psi(x,t)\,dx\,dy\end{aligned} \tag{7-97}$$

Problem 7-16 Show that

$$\left\langle\chi\left|x_l\frac{m(x_{k+1}-x_k)}{\epsilon}\right|\psi\right\rangle=\int\chi^*(x,t)xK(x,t;y,s)\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right)\psi(y,s)\,dy\,dx \tag{7-98}$$

if $t_l=t$ and $t_k=s$, provided $t_l>t_k$. What happens if $t_l<t_k$?

Notice that the square of the momentum p^2 corresponds to pp, or two *successive* velocities times mass multiplied together (as in Prob. 7-15). It does *not* correspond to the simple square of velocity at one time, $\langle\chi|m^2(x_{k+1} - x_k)^2/\epsilon^2|\psi\rangle$; for that goes to infinity as $m\hbar/i\epsilon$ when $\epsilon \to 0$, as we have seen in Sec. 7-3, particularly in Eq. (7-49). The difference between this expression $m\hbar/i\epsilon$ and the left-hand side of Eq. (7-97) is in fact p^2 in the limit. That is,

$$\left\langle \chi \left| \frac{m^2(x_{k+1} - x_k)^2}{\epsilon^2} \right| \psi \right\rangle = \frac{m\hbar}{i\epsilon} \langle\chi|1|\psi\rangle + \left\langle \chi \left| m \frac{x_{k+1} - x_k}{\epsilon} m \frac{x_k - x_{k-1}}{\epsilon} \right| \psi \right\rangle \tag{7-99}$$

Problem 7-17 Prove this, using Eq. (7-40) with

$$F = \frac{m}{\epsilon}(x_{k+1} - x_k)$$

7-6 THE PERTURBATION SERIES FOR A VECTOR POTENTIAL

The singular behavior of the transition element of the square of the velocity, as shown in Eq. (7-49), has as a consequence the fact that many expressions involving velocity must be translated with care. For example, the lagrangian of a particle in an electromagnetic field is

$$L = \frac{m}{2}|\dot{\mathbf{r}}|^2 + eV(\mathbf{x},t) - \frac{e}{c}\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r},t) \tag{7-100}$$

Let us take $V = 0$ and ask for the effect of the vector potential $\mathbf{A}$ considered as a perturbation. That is, with $S_0 = (m/2)\int|\dot{\mathbf{r}}|^2\,dt$, $\sigma = -(e/c)\int\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r},t)\,dt$, we develop a series for use in a perturbation treatment and solve for the resulting transition elements. Thus

$$\langle e^{i\sigma/\hbar}\rangle_{S_0} = \langle 1\rangle_{S_0} + \frac{i}{\hbar}\langle\sigma\rangle_{S_0} - \frac{1}{2\hbar^2}\langle\sigma^2\rangle_{S_0} + \cdots \tag{7-101}$$

The first-order term is $-ie/\hbar c$ times the expression

$$\langle\textstyle\int\dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r},t)\,dt\rangle \tag{7-102}$$

We wish to translate this to operator notation. In defining σ for a discontinuous path (a series of steps of time length ϵ) we might at first expect to write either

$$\sigma = -\frac{e}{c}\sum_k (\mathbf{r}_{k+1} - \mathbf{r}_k) \cdot \mathbf{A}(\mathbf{r}_k,t) \tag{7-103}$$

or

$$\sigma = -\frac{e}{c}\sum_k (\mathbf{r}_{k+1} - \mathbf{r}_k) \cdot \mathbf{A}(\mathbf{r}_{k+1},t_{k+1}) \tag{7-104}$$

Either one, in the limit for a continuous path, gives the integral for σ. But if we look at a particular component of **A**, say, A_x, we find that $A_x(\mathbf{r}_{k+1},t_{k+1})$ differs from $A_x(\mathbf{r}_k,t_k)$ by approximately

$$(\mathbf{r}_{k+1} - \mathbf{r}_k) \cdot \nabla A_x + \epsilon \frac{\partial A_x}{\partial t} \tag{7-105}$$

which, when multiplied by $\mathbf{r}_{k+1} - \mathbf{r}_k$ again, might be expected to be of second order for each k (leading to a term only of order ϵ when k is summed). But our paths are not continuous and the transition element of the mean square of $x_{k+1} - x_k$ is of first order. In fact (see Prob. 7-6),

$$(x_{k+1} - x_k)^2 \approx -\frac{\hbar\epsilon}{mi}$$

$$(x_{k+1} - x_k)(y_{k+1} - y_k) \approx 0$$

$$(y_{k+1} - y_k)^2 \approx -\frac{\hbar\epsilon}{mi}$$

etc., to first order in ϵ. Hence Eq. (7-103) differs from Eq. (7-104) by approximately

$$\sum_k \frac{\hbar\epsilon}{mi} \nabla \cdot \mathbf{A}(\mathbf{r}_k,t_k) = \frac{\hbar}{mi}\int \nabla \cdot \mathbf{A}\, dt \tag{7-106}$$

a zero-order term. So it is imperative to decide which form is correct.

The general answer to such a question was given in Chap. 2. There the rule given was that S is replaced by $\sum_k S_{cl}(x_{k+1},t_{k+1};x_k,t_k)$, where S_{cl} is the classical action to go from one point to a neighboring point. It is not necessary to calculate this action exactly, but only sufficiently closely to resolve ambiguities. Equations (7-103) and (7-104) are not sufficiently closely calculated for this purpose, but the classical action for a short interval is very close to

$$S_{cl}[k+1, k] = \frac{m|\mathbf{r}_{k+1} - \mathbf{r}_k|^2}{2\epsilon} + \frac{1}{2}[\mathbf{A}(\mathbf{r}_{k+1},t_{k+1}) + \mathbf{A}(r_k,t_k)] \cdot (\mathbf{r}_{k+1} - r_k) \tag{7-107}$$

Therefore, the correct expression for σ is the average of Eqs. (7-103) and (7-104), so that the transition element of Eq. (7-99) is

$$\left\langle \sum_k (\mathbf{r}_{k+1} - \mathbf{r}_k) \cdot \tfrac{1}{2}[\mathbf{A}(\mathbf{r}_{k+1},t_{k+1}) + \mathbf{A}(\mathbf{r}_k,t_k)] \right\rangle \tag{7-108}$$

Leaving the sum on k for later evaluation (as an integral over time) the result is expressed as the operator $(1/2m)(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p})$ (see Prob. 7-12).

That is, in an electromagnetic potential, the first-order term in the perturbation expansion has the same form as the first-order term given in Eq. (6-11), but the quantity V is replaced by the operator $-e/(2cm)(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p})$.

This conclusion is not true in the second order. The second-order term requires our finding

$$\frac{1}{2}\left(\frac{e}{\hbar c}\right)^2 \langle [\textstyle\int \dot{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r},t)\,dt]^2 \rangle = \frac{1}{2}\left(\frac{e}{\hbar c}\right)^2 \sum_k \sum_l \left\langle \left\{ [\mathbf{r}_{k+1} - \mathbf{r}_k] \cdot \frac{1}{2}[\mathbf{A}(\mathbf{r}_{k+1},t_{k+1}) + \mathbf{A}(\mathbf{r}_k,t_k)] \right\} \left\{ [\mathbf{r}_{l+1} - \mathbf{r}_l] \cdot \frac{1}{2}[\mathbf{A}(\mathbf{r}_{l+1},t_{l+1}) + \mathbf{A}(\mathbf{r}_l,t_l)] \right\} \right\rangle \tag{7-109}$$

Nothing special happens for the terms $k \neq l$, and we obtain in fact precisely the second-order term expected by comparison to Eq. (6-13) with V replaced by the operator $-e/(2cm)(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p})$. But when $l = k$, the coincidence of the two velocities gives a new term. In view of Eq. (7-49) and Prob. 7-6 we get an additional quantity

$$\frac{e^2}{2c^2}\left\langle + \frac{i\hbar\epsilon}{m} \sum_k \left\{ \frac{1}{2}[\mathbf{A}(\mathbf{r}_{k+1},t_{k+1}) + \mathbf{A}(\mathbf{r}_k,t_k)] \right\}^2 \right\rangle \tag{7-110}$$

which is equivalent to $ie^2/2mc^2\int[\mathbf{A}(\mathbf{r},t) \cdot \mathbf{A}(\mathbf{r},t)]\,dt$ and has the same effect as the first-order action of a potential $(e^2/2mc^2)\mathbf{A} \cdot \mathbf{A}$.

Thus the perturbation expansion for the action of a vector potential has the same form as Eq. (6-17). The potential V is replaced by the operator $-e/2mc(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + e^2/2mc^2\mathbf{A} \cdot \mathbf{A}$. We have shown it to second order in A, but a little consideration shows it is true to any order.

The hamiltonian for a particle in a vector potential $\mathbf{A}$ is

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \cdot \left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right) \tag{7-111}$$

It differs from that of a free particle [which is $(1/m)\mathbf{p} \cdot \mathbf{p}$] by just this operator $-e/2mc(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + e^2/2mc^2\mathbf{A} \cdot \mathbf{A}$. This is a much easier way to arrive at the result we have just obtained.

7-7 THE HAMILTONIAN

Using what we have so far derived, it would be very easy to write down the transition amplitude for the hamiltonian. We take the transition amplitude for the square of the momentum, divide it by $2m$, and add the transition amplitude for the potential. In this way the hamiltonian itself at the time t_k could be written as

$$H_k = \frac{m}{2}\left(\frac{x_{k+1} - x_k}{\epsilon}\right)\left(\frac{x_k - x_{k-1}}{\epsilon}\right) + V(x_k) \tag{7-112}$$

while in operator form we have the transition element of the hamiltonian as

$$\langle\chi|H_k|\psi\rangle = \int_{-\infty}^{\infty} \chi^*\left[\frac{p^2}{2m} + V(x)\right]\psi\,dx = \int_{-\infty}^{\infty} \chi^* H\psi\,dx \tag{7-113}$$

Although this method for defining the transition amplitude for the hamiltonian gives a perfectly correct result, it is somewhat artificial, since it does not exhibit the important relationship between the hamiltonian and time. Therefore, we shall next consider an alternative definition of this transition element based upon an investigation of the changes made in a state when it is displaced in time. This approach will also enable us to define H_k given only the form of S, no matter how complicated.

To carry out this investigation, we break up the time axis into infinitesimal intervals, just as we did in defining path integrals. Now, however, it is important to point out that the subdivision of time into equal intervals is not necessary. Clearly, any subdivision of time into equal intervals is not necessary. Any subdivision into instants t_i will be satisfactory; the process of taking limits is characterized by having the largest spacing $t_{i+1} - t_i$ approach zero.

For simplicity, our system will consist of a single particle moving in one dimension. The action is represented by the sum

$$S = \sum_i S[x_{i+1},t_{i+1};x_i,t_i] \tag{7-114}$$

where

$$S[x_{i+1},t_{i+1};x_i,t_i] = \int_{t_i}^{t_{i+1}} L[\dot{x}(t),x(t)]\,dt \tag{7-115}$$

The integral in this expression is taken along the classical path between x_i at t_i and x_{i+1} at t_{i+1}. For our one-dimensional example we can write, with sufficient accuracy,

$$S[x_{i+1},t_{i+1};x_i,t_i] = \left[\frac{m}{2}\left(\frac{x_{i+1}-x_i}{t_{i+1}-t_i}\right)^2 - V(x_{i+1})\right](t_{i+1}-t_i) \tag{7-116}$$

The normalization constant associated with an integral over dx_i at the time t_i is the same one we have used before, namely,

$$A = \left[\frac{2\pi\hbar i(t_{i+1}-t_i)}{m}\right]^{1/2} \tag{7-117}$$

The relation of H to the change in a state with displacement in time can now be studied. Consider a state $\psi(t)$ defined by a space-time region R. Now imagine that at the same time t we consider another state $\psi_\delta(t)$, defined by another region R_δ. Suppose the region R_δ is exactly the same as R except that it is earlier by a time δ, that is, displaced bodily toward the past by a time δ. All the apparatus required to prepare the system for R_δ is identical to that for R but is operated a time δ sooner. If L depends explicitly on time, it too is to be displaced; i.e., the state ψ_δ is obtained from the L used for the state ψ except that in writing L_δ we use as the time variable $t + \delta$.

Now we ask: How does the state ψ_δ differ from ψ? In any measurement the chance of finding the system in some fixed region R' is different depending on whether the region of origin was R or R_δ. Consider the change in transition amplitude $\langle\chi|1|\psi_\delta\rangle$ produced by the shift in time δ. We can consider this shift as effected by decreasing all values of t_i by δ for $i \leq k$ and leaving all t_i fixed for $i > k$.

If the reader looks ahead at this point, it may occur to him that we are headed for trouble. Clearly, it is our intention eventually to take a limit as all infinitesimal time intervals are decreased to zero. However, with the present setup, at least one time interval $t_{k+1} - t_k$ has a lower bound, so that it cannot be indefinitely decreased. This difficulty could be straightened out by assuming the time shift δ to be itself a function of time. We can imagine that it is turned on smoothly before $t = t_k$ and turned off smoothly after $t = t_k$. Then keeping the time variation of δ fixed, we can let all time intervals proceed smoothly to zero, including $t_{k+1} - t_k$. We would then investigate the first-order effect of the time shift by letting the magnitude of δ approach zero. The result obtained by this more rigorous process is essentially the same as that of the procedure we are using in our present example.

Returning now to our investigation of the effect of the time shift we see that the action $S[x_{i+1},t_{i+1};x_i,t_i]$ as defined by Eq. (7-115) will not be changed so long as both t_{i+1} and t_i are changed by the same amount. On the other hand, $S[x_{k+1},t_{k+1};x_k,t_k]$ is changed to $S[x_{k+1},t_{k+1};x_k,t_k - \delta]$.

Furthermore, the constant A associated with the integration over dx_k is also altered and becomes

$$A_k = \left[\frac{2\pi\hbar i(t_{k+1} - t_k + \delta)}{m}\right]^{1/2} \tag{7-118}$$

We use Eq. (7-2) to define the transition amplitude. Keeping in mind that the path integral depends on both the action S and the normalizing constant A (both of which are altered by our time shift) we can write the change in the transition amplitude to first order in δ as

$$\langle\chi|1|\psi\rangle - \langle\chi|1|\psi_\delta\rangle = \left\langle\chi\left|\frac{\partial S[x_{k+1},t_{k+1};x_k,t_k]}{\partial t_k} + \frac{\hbar}{2i(t_{k+1}-t_k)}\right|\psi\right\rangle\frac{i\delta}{\hbar} \tag{7-119}$$

the second term coming from the change in A. We wish to define the functional corresponding to the hamiltonian in quantum mechanics as

$$H_k = \frac{\partial S[x_{k+1},t_{k+1};x_k,t_k]}{\partial t_k} + \frac{\hbar}{2i(t_{k+1}-t_k)} \tag{7-120}$$

The first term on the right-hand side of this last equation is the definition of the classical hamiltonian. The second term is necessary in the quantum-mechanical definition in order to keep H_k finite as the time interval $t_{k+1} - t_k$ goes to zero. This last term is a consequence of the change in the normalizing constant A due to the time shift δ.

Applying this result to the specific one-dimensional example indicated by Eq. (7-116), we can write the operator H_k as

$$\begin{aligned} H_k &= \frac{m}{2}\left(\frac{x_{k+1}-x_k}{t_{k+1}-t_k}\right)^2 + \frac{\hbar}{2i(t_{k+1}-t_k)} + V(x_{k+1}) \\ &= \frac{m}{2}\left(\frac{x_{k+1}-x_k}{t_{k+1}-t_k}\right)\left(\frac{x_k-x_{k-1}}{t_k-t_{k-1}}\right) + V(x_k) \end{aligned} \tag{7-121}$$

The second of these equations is based upon the results obtained in Eq. (7-54). By writing the product of velocities as the product of two successive velocities, we can do away with the apparently extraneous term $\hbar/[2i(t_{k+1} - t_k)]$.

Using the relation $t_\delta = t - \delta$ for all values of $t < t_k$, we have

$$\psi(t) = \psi(t_\delta) + \delta\frac{\partial\psi}{\partial t} = \psi_\delta + \delta\frac{\partial\psi}{\partial t} \tag{7-122}$$

connecting the function ψ defined in the two regions R and R_δ. Thus the cycle of relations connecting operators to the Schrödinger equation and to path integrals can be closed with the result obtained by combining Eqs. (7-119), (7-120), and (7-122).

$$\delta\left\langle\chi|1|\frac{\partial\psi}{\partial t}\right\rangle = \frac{i}{\hbar}\,\delta\langle\chi|H_k|\psi\rangle \tag{7-123}$$

which leads us back again to the Schrödinger equation

$$\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = H\psi \tag{7-124}$$

For arbitrarily complicated actions we can find an expression for the hamiltonian (i.e., a functional corresponding to the energy) by asking for the first-order change in the transition element $\langle\chi|1|\psi\rangle$ when all times previous to t are shifted by $-\delta$ and writing this change as $\delta\langle\chi|H(t)|\psi\rangle$.

8

Harmonic Oscillators

THE problem of the harmonic oscillator is perhaps the simplest in quantum mechanics. We have solved it completely in noting that[1] the kernel for the motion of a harmonic oscillator is

$$K(x_b,T;x_a,0) = \left(\frac{m\omega}{2\pi i\hbar \sin \omega T}\right)^{1/2} \times \exp\left\{\frac{im\omega}{2\hbar \sin \omega T}[(x_a^2 + x_b^2)\cos \omega T - 2x_ax_b]\right\} \quad (8\text{-}1)$$

If we are to make full use of this, we should look at all sorts of problems which involve harmonic oscillators, either exactly or approximately. It is the purpose of this chapter to describe several such problems, both those involving single oscillators and those involving systems of interacting harmonic oscillators. We could carry this program to extremes and include all kinds of classical vibration problems, plates, rods, etc., but such systems are so large that it would be a waste of time to analyze the quantum-mechanical corrections. Instead, it would be better to look at systems on the atomic scale. For example, we might analyze the oscillations of the molecule CO. In so doing, we find that the potential energy between the carbon and oxygen atoms is not exactly quadratic. Nevertheless, for the lower-energy state the potential is so close to quadratic that a pure harmonic oscillator treatment is a good approximation for many purposes.

In a much more complicated polyatomic molecule, when the excitation energy is not too high, the travel of the atoms is small compared with their spacing. In this case again the potential energy is very nearly a quadratic function of the coordinates. Thus the system is approximately equivalent to a set of coupled harmonic oscillators. A solid crystal is, from one point of view, a polyatomic molecule of great size. As such it is a vast array of interacting harmonic oscillators.

As another example we can consider the electromagnetic field in a cavity. Classically, there are several patterns of standing waves, or modes, in which the field can vibrate harmonically with a definite frequency. In quantum mechanics, each of these modes constitutes a quantum oscillator.

8-1 THE SIMPLE HARMONIC OSCILLATOR

Solution from the Schrödinger Equation. In this section we shall develop a number of relations describing the simple one-dimensional harmonic oscillator. We shall begin with the language of the Schrödinger equation. Problem 2-2 gave the lagrangian describing a one-dimensional harmonic oscillator as

---

[1] Cf. Prob. 3-8.

$$L = \frac{m}{2} (\dot{x}^2 - \omega^2 x^2) \tag{8-2}$$

The corresponding hamiltonian, which we use in the present treatment, is

$$H = \frac{p^2}{2m} + \frac{m}{2} \omega^2 x^2 \tag{8-3}$$

The wave equation is then

$$\begin{aligned} -\frac{\hbar}{i} \frac{\partial \psi}{\partial t} &= H\psi \\ &= \left(\frac{p^2}{2m} + \frac{m}{2} \omega^2 x^2 \right) \psi \end{aligned} \tag{8-4}$$

Since the hamiltonian is independent of time, the wave equation is easily separated, and it yields wave functions of steady states of definite energy E_n. The time-dependent part is proportional to exp $(iE_n t/\hbar)$.

Recalling that the momentum operator p corresponds to differentiation with respect to x (cf. Sec. 7-5), we can write the Schrödinger equation for the spatial part of the wave function as

$$H\phi_n = \frac{\hbar^2}{2m} \frac{\partial^2 \phi_n}{\partial x^2} + \frac{m\omega^2 x^2}{2} \phi_n = E_n \phi_n \tag{8-5}$$

This equation is easily solved. The result is given in many books on quantum mechanics.[1] The eigenvalues for the energy are

$$E_n = \hbar\omega(n + \tfrac{1}{2}) \tag{8-6}$$

where n is an integer, 0, 1, 2, ... The eigenfunctions ϕ_n are

$$\phi_n = (2^n n!) \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} H_n\left(x \sqrt{\frac{m\omega}{\hbar}} \right) e^{-(m\omega x^2/2\hbar)} \tag{8-7}$$

where the functions H_n are the Hermite polynomials

$$H_0(y) = 1$$

$$H_1(y) = 2y$$

$$H_2(y) = 4y^2 - 2$$

. . . .

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-y^2} \tag{8-8}$$

---

[1] L. I. Schiff, "Quantum Mechanics," 2d ed., McGraw-Hill Book Company, New York, 1955.

The Hermite polynomials are best defined by their generating function

$$e^{-t^2+2ty} = \sum_{n=0}^{\infty} H_n(y) \frac{t^n}{n!} \tag{8-9}$$

We can obtain these results in another manner. The functions ϕ_n have been obtained by solving a differential equation, namely, the time-independent case. However, we already have a solution for the time-dependent case. From this solution we should be able to derive these functions directly. It is instructive to carry out this derivation to illustrate some of the formulas which have been derived in earlier chapters.

Solution from the Kernel. We have worked out the kernel describing the motion of an oscillator in Prob. 3-8. On the other hand, we know from Eq. (4-59) that this kernel can be expanded in exponential functions of time multiplied by products of energy eigenfunctions. That is,

$$\left(\frac{m\omega}{2\pi i\hbar \sin \omega T}\right)^{1/2} \exp\left(\frac{im\omega}{2\hbar \sin \omega T}\right) [(x_1^2 + x_2^2) \cos \omega T - 2x_1x_2]$$

$$= \sum_{n=0}^{\infty} e^{-(i/\hbar)E_nT} \phi_n(x_2)\phi_n^*(x_1) \tag{8-10}$$

Using the relations

$$\begin{aligned} i \sin \omega T &= \tfrac{1}{2}e^{i\omega T}(1 - e^{-2i\omega T}) \\ \cos \omega T &= \tfrac{1}{2}e^{i\omega T}(1 + e^{-2i\omega T}) \end{aligned} \tag{8-11}$$

we can write the left-hand side of Eq. (8-10) as

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(i\omega T/2)}(1 - e^{-2i\omega T})^{-1/2}$$

$$\times \exp\left\{-\frac{m\omega}{2\hbar}\left[(x_1^2 + x_2^2)\left(\frac{1 + e^{2i\omega T}}{1 - e^{2i\omega T}}\right) - \frac{4x_1x_2e^{-i\omega T}}{1 - e^{2i\omega T}}\right]\right\} \tag{8-12}$$

We can obtain a series having the form of the right-hand side of Eq. (8-10) if we expand Eq. (8-12) in successive powers of $e^{-i\omega T}$. Because of the initial factor $e^{-i\omega T/2}$, it is clear that all terms in the expansion will be of the form $e^{-i\omega T/2}e^{-in\omega T}$ for $n = 0, 1, 2, \ldots$. This means the energy levels are given by

$$E_n = \hbar\omega(n + \tfrac{1}{2}) \tag{8-13}$$

To find the wave functions, we shall have to carry out the expansion completely. We shall illustrate the method by going only as far as $n = 2$. Expanding the left-hand side of Eq. (8-10) to this order we have

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-i\omega T/2}\left(1 + \tfrac{1}{2}e^{-2i\omega T} + \cdots\right) \exp\left[-\frac{m\omega}{2\hbar}(x_1^2 + x_2^2) - \frac{m\omega}{\hbar}(x_1^2 + x_2^2)(e^{-2i\omega T} + \cdots) + \frac{2m\omega}{\hbar}x_1x_2e^{-i\omega T} + \cdots\right] \tag{8-14}$$

or

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(m\omega/2\hbar)(x_1^2+x_2^2)}e^{-i\omega T/2}\left(1 + \tfrac{1}{2}e^{-2i\omega T}\right) \times\left[1 + \frac{2m\omega}{\hbar}x_1x_2e^{-i\omega T} + \frac{4m^2\omega^2}{2\hbar^2}x_1^2x_2^2e^{-2i\omega T} - \frac{m\omega}{\hbar}(x_1^2 + x_2^2)e^{-2i\omega T}\cdots\right] \tag{8-15}$$

From this we can pick out the coefficient of the lowest term. It is

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(m\omega/2\hbar)(x_1^2+x_2^2)}e^{-(i\omega T/2)} = e^{-(i/\hbar)E_0T}\phi_0(x_2)\phi_0^*(x_1) \tag{8-16}$$

This means that $E_0 = \hbar\omega/2$ and

$$\phi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-(m\omega x^2/2\hbar)} \tag{8-17}$$

We have chosen ϕ_0 to be real. We could make it complex by including the factor $e^{i\delta}$, where δ is a constant; however, it would make no difference to any physical result.

The next-order term in the expansion is

$$e^{-i\omega T/2}e^{-i\omega T}\frac{m\omega}{\pi\hbar}e^{-(m\omega/2\hbar)(x_1^2+x_2^2)}\frac{2m\omega}{\hbar}x_1x_2 = e^{-(i/\hbar)E_1T}\phi_1(x_2)\phi_1^*(x_1) \tag{8-18}$$

which implies that $E_1 = \tfrac{3}{2}\hbar\omega$, and

$$\phi_1(x) = \frac{2m\omega}{\hbar}x\phi_0(x) \tag{8-19}$$

The next term corresponds to $E_2 = \tfrac{5}{2}\omega\hbar$. The part of the term depending on x_1 and x_2 is

$$\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} e^{-(m\omega/2\hbar)(x_1^2+x_2^2)}\left[\frac{2m^2\omega^2}{\hbar^2}x_1^2x_2^2 - \frac{m\omega}{\hbar}(x_1^2 + x_2^2)\right] \tag{8-20}$$

This must be the same as $\phi_2(x_2)\phi_2^*(x_1)$. Since the expression in the brackets can be rewritten as

$$\frac{1}{2}\left(\frac{2m\omega}{\hbar}x_1{}^2 - 1\right)\left(\frac{2m\omega}{\hbar}x_2{}^2 - 1\right) \tag{8-21}$$

we find

$$\phi_2(x) = \frac{1}{\sqrt{2}}\left(\frac{2m\omega}{\hbar}x^2 - 1\right)\phi_0(x) \tag{8-22}$$

These results can be compared with those obtained from the solution of the wave equation, given by Eqs. (8-7) and (8-8).

All of the wave functions may be obtained in this manner. However, it is a difficult algebraic problem to get the general form for ϕ_n directly from this expansion. A less direct way is illustrated in the problem.

Problem 8-1 Note that the amplitude to go from any state $f(x)$ to another state $g(x)$ is the transition amplitude $\langle g|1|f\rangle$ as defined in Eq. (7-1).

Suppose $f(x)$ and $g(x)$ are expanded in terms of the orthogonal functions $\phi_n(x)$, the solutions to the wave equation associated with the kernel $K(2,1)$, as discussed in Sec. 4-2. Thus

$$f(x) = \Sigma f_n\phi_n(x) \qquad g(x) = \Sigma g_n\phi_n(x) \tag{8-23}$$

Using the coefficients f_n and g_n and Eq. (4-59), show that the transition amplitude can be written as

$$\int\int g^*(x_2)K(x_2,T;x_1,0)f(x_1)\,dx_1\,dx_2 = \Sigma g_n^* f_n e^{-(i/\hbar)E_nT} \tag{8-24}$$

Next, suppose we choose a special pair of functions f and g for which the expansion on the right-hand side of Eq. (8-24) is simple. Then after obtaining the functions f_n we could get some information about the wave function ϕ_n from the expansions of Eq. (8-23). Suppose we choose the functions f and g in the following way,

$$f(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-(m\omega/2\hbar)(x-a)^2} \tag{8-25}$$

$$g(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-(m\omega/2\hbar)(x-b)^2} \tag{8-26}$$

These functions represent gaussian distributions centered about a and b, respectively. We shall call $f_n = f_n(a)$ and $g_n = f_n(b)$. Determine the transition amplitude $\langle f|1|g\rangle$, where f and g are given by Eqs. (8-25) and (8-26), respectively, and the kernel is that for a harmonic oscillator, Eq. (8-1). Perform the integrals in Eq. (8-24) to get

$$\exp\left[-\frac{i\omega T}{2} - \frac{m\omega}{4\hbar}(a^2 + b^2 - 2abe^{-i\omega T})\right] = \sum_n f_n(a) f_n^*(b) e^{-(i/\hbar)E_nT} \tag{8-27}$$

From this result show that $E_n = \hbar\omega(n + \frac{1}{2})$ and that

$$f_n(a) = \left(\frac{m\omega}{2\hbar}\right)^{n/2} \frac{a^n}{\sqrt{n!}} \exp\left(-\frac{m\omega a^2}{4\hbar}\right) \tag{8-28}$$

Use this result in Eq. (8-24) and write for ϕ_n the form given by Eq. (8-7) considering the $H_n(x)$ still unknown. From this derive the generating function of Eq. (8-9) for these functions $H_n(x)$.

8-2 THE POLYATOMIC MOLECULE

In the preceding section we derived the wave functions and energy levels which describe the simple harmonic oscillator. In this section we begin our investigation of systems of interacting oscillators with the study of polyatomic molecules. We begin the analysis by assigning coordinates describing the position of each atom in the molecule. The position of any particular atom a will be given by the three cartesian coordinates x_a, y_a, and z_a, whose origin lies at the equilibrium position for the atom. If the mass of the atom is m_a, the kinetic energy of the whole molecule is given by

$$\sum_a \frac{1}{2} m_a(\dot{x}_a^2 + \dot{y}_a^2 + \dot{z}_a^2) \tag{8-29}$$

where the summation is carried out over all atoms in the molecule.

It will be more convenient for this general discussion to avoid the vector aspects of this description by making the following modification. We suppose there are N atoms in the molecule. We shall define $n = 3N$ cartesian coordinates in the following way,

$$\begin{aligned} &q_1 = \sqrt{m_a}\,x_a \qquad q_2 = \sqrt{m_a}\,y_a \qquad q_3 = \sqrt{m_a}\,z_a \\ &q_4 = \sqrt{m_b}\,x_b \qquad q_5 = \sqrt{m_b}\,y_b \qquad \cdots \end{aligned} \tag{8-30}$$

In terms of these new coordinates the kinetic energy has the form

$$\text{K.E.} = \sum_{j=1}^{n} \frac{1}{2} \dot{q}_j^2 \tag{8-31}$$

The potential energy is the function $V(q_1,q_2, \ldots)$ of all the displacements q_j. We can expand V in a Taylor series around the equilibrium position $q_j = 0$. Thus

$$V(q_1,q_2, \ldots ,q_n) = V(0,0, \ldots ,0) + \sum_{j=1}^{n} q_j V_j(0,0, \ldots ,0) + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} q_j q_k V_{jk}(0,0, \ldots ,0) + \cdots \tag{8-32}$$

where

$$V_j = \frac{\partial V}{\partial q_j} \qquad V_{jk} = \frac{\partial^2 V}{\partial q_j \, \partial q_k} \tag{8-33}$$

The first term is the potential energy at equilibrium. It is a constant independent of q. We shall assign the value zero to this constant potential and measure all other potential energies in a manner consistent with that assignment. Thus the first term can be omitted from the expansion. The factor $V_j(0,0, \ldots ,0)$ appears in the next term. This is the potential gradient or force associated with the coordinate q_j and evaluated in equilibrium position. This factor is therefore zero, and the term can be omitted from the series. To put this another way, since equilibrium corresponds to a minimum of potential energy, there must be a vanishing first-order change for displacements around the equilibrium position.

The factors $V_{jk}(0,0, \ldots)$ appearing in the next term comprise a set of constants whose value depends on the structure of the molecule. Call these constants v_{jk}. Now suppose we neglect all higher-order terms. Thus we make the approximation that the potential energy involves the coordinates only quadratically. Even if the potential is not a pure quadratic function of the coordinates, our approximation will be valid for small displacements. It is by this approximation that we represent our molecule as a system of harmonic oscillators.

Combining Eqs. (8-31) and (8-32), we can write the lagrangian as

$$L = \frac{1}{2} \sum_{j=1}^{n} \dot{q}_j^2 - \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} q_j q_k v_{jk} \tag{8-34}$$

Next, we introduce this lagrangian into the path integral which defines the kernel describing the motion of the atoms in the molecule,

$$K = \iint \cdots \int \left(\exp \left\{ \frac{i}{2\hbar} \left[\int \sum_{j=1}^{n} \dot{q}_j^2(t)\, dt \right.\right.\right.$$
$$\left.\left.\left. - \sum_{j=1}^{n} \sum_{k=1}^{n} v_{jk} \int q_j(t) q_k(t)\, dt \right] \right\} \right) \mathfrak{D}q_1(t)\, \mathfrak{D}q_2(t) \cdots \mathfrak{D}q_n(t) \qquad (8\text{-}35)$$

All of these path integrals are gaussian, and thus they can be solved by the methods discussed in Sec. 3-5. To carry out that solution, we shall have to find those paths $q_j(t)$ which give a stationary value for the action integral. Variation with respect to each q_j gives these paths as solutions of

$$\ddot{q}_j = - \sum_{k=1}^{n} v_{jk} q_k \qquad (8\text{-}36)$$

This last equation says that the force on any single atom in a particular direction is some linear combination of the displacements of all the atoms.

Such systems of interacting oscillators have been analyzed to a great extent from a classical point of view. Since in many problems of quantum mechanics we obtain the classical action as the first step in solving the kernels, all of this classical work will be of great value to us. One important result of the classical analysis is the following. There are special ways to distort the molecule so that, as time goes on, the motion is of the simple periodic sinusoidal type. The pattern of distortions remains the same, and only the amount of the distortion varies sinusoidally with time. Different patterns of distortion, or, as we say, different modes, correspond in general to different frequencies. There may be some with zero frequency, and some groups of modes may all have the same frequency. The important fact is this: Any small displacement motion of the molecule can be built up as a linear combination of such modes.

If there are N atoms in the molecule, then the molecule has $n = 3N$ modes of motion. Thus, for example, the molecule CO_2 has nine modes, as shown by Fig. 8-1, where the motion of each atom is indicated by an arrow. Only modes 1 to 4 are periodic (i.e., have a nonzero frequency) and the direction of motion during the first half-cycle is indicated. For the second half-cycle, reverse all arrows.

We shall next derive the mathematical description of the modes. This derivation is, of course, part of classical physics, rather than

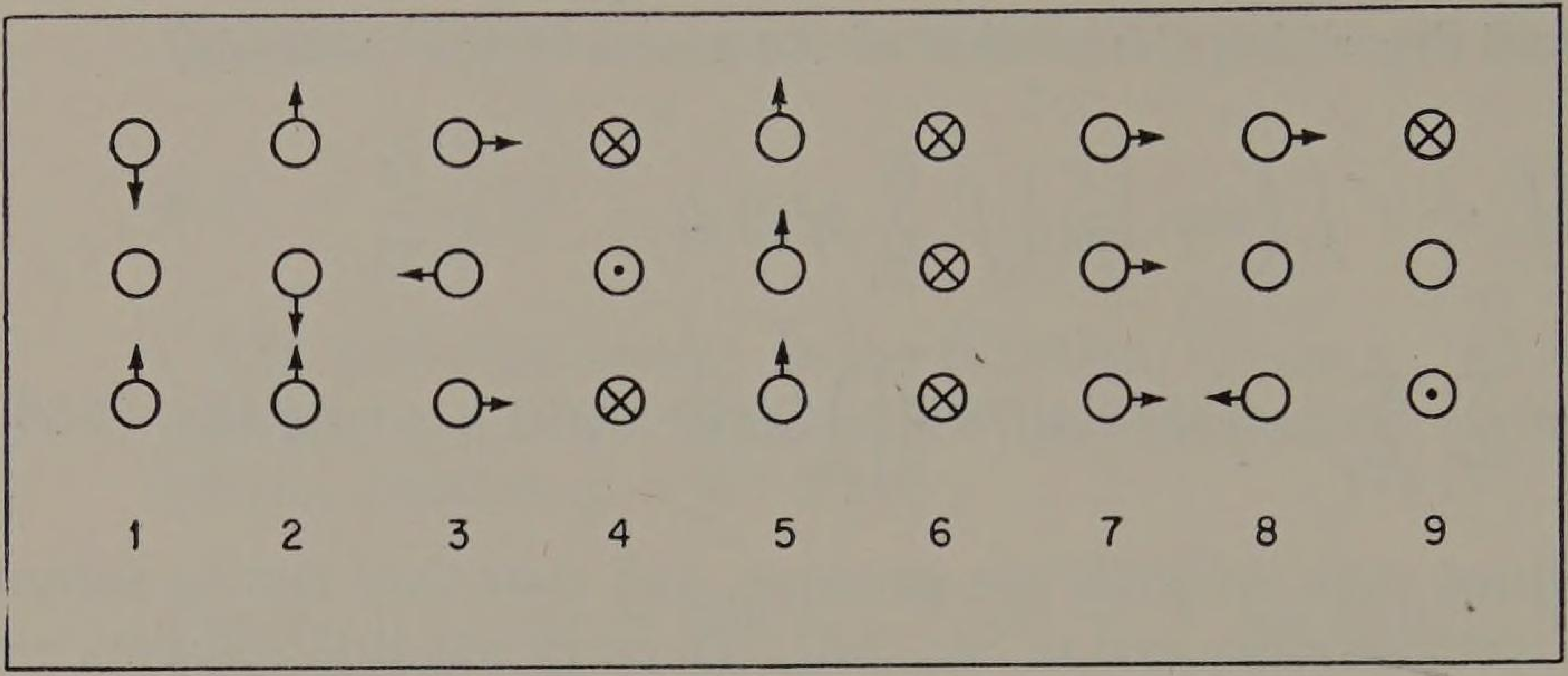


Fig. 8-1 Normal modes of the CO_2 molecule. The symbol ⊙ means motion out of the plane of the paper, and ⊗ means motion into the plane. Modes 1 to 4 are periodic; modes 5 to 7 are continuous translations; and modes 8 and 9 are continuous rotations.

quantum mechanics. Consider a particular mode of frequency ω. All of the coordinates q_j move together and at the same frequency. There must be some special set of initial displacements a_j, different for each mode, such that if all initial velocities are zero, the subsequent motion of any coordinate can be written as

$$q_j = a_j \cos \omega t \tag{8-37}$$

Substituting this equation into Eq. (8-36) gives

$$\omega^2 a_j = \sum_{k=1}^{n} v_{jk} a_k \tag{8-38}$$

The last equation is actually a set of n equations for the n unknowns a_j. Since it is homogeneous, it has a solution only if the determinant of coefficients vanishes. Thus we require

$$\begin{bmatrix} (\omega^2 - v_{11}) & -v_{12} & \cdots & -v_{1n} \\ -v_{21} & (\omega^2 - v_{22}) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ -v_{n1} & \cdots & \cdots & (\omega^2 - v_{nn}) \end{bmatrix} = 0 \tag{8-39}$$

This equation has n solutions for ω^2. For a particular solution, say ω_α, we can get solutions for the set of equations (8-38). We shall call these $a_{j\alpha}$. The sizes of the constants $a_{j\alpha}$ are determined relative to each other, but the absolute magnitude of the whole set is arbitrary. We shall choose this magnitude so that

$$\sum_{j=1}^{n} a_{j\alpha}^2 = 1 \tag{8-40}$$

We can repeat this process for all of the n modes, $\alpha = 1, 2, \ldots, n$. We determine n values of ω_α, and for each value of α we obtain n solutions for the j constants $a_{j\alpha}$. Any possible motion of the system is a linear combination of these modes. We can write an expression for a general type of motion as

$$q_j = \sum_{\alpha=1}^{n} C_\alpha a_{j\alpha} \cos(\omega_\alpha t + \delta_\alpha) \tag{8-41}$$

Here the constant of amplitude C_α and the constant phase δ_α depend on the initial conditions. That such an expression does represent the motion of the system is easily verified by substituting Eq. (8-41) into Eq. (8-36).

It is convenient to use the complex notation for Eq. (8-41). That is,

$$q_j = \sum_{\alpha=1}^{n} C_\alpha a_{j\alpha} e^{i\omega_\alpha t} e^{i\delta_\alpha} = \sum_{\alpha=1}^{n} c_\alpha a_{j\alpha} e^{i\omega_\alpha t} \tag{8-42}$$

where only the real part of the expression has physical meaning. The complex constants c_α depend on the initial conditions, and they can be determined as follows. If the initial positions and velocities are $q_j(0)$ and $\dot{q}_j(0)$, respectively, we have

$$\begin{aligned} q_j(0) &= \text{Real Part} \sum_{\alpha=1}^{n} c_\alpha a_{j\alpha} = \sum_{\alpha=1}^{n} (\text{R.P. } c_\alpha) a_{j\alpha} \\ \dot{q}_j(0) &= \text{R.P.} \sum_{\alpha=1}^{n} i c_\alpha a_{j\alpha} \omega_\alpha = \sum_{\alpha=1}^{n} [-(\text{Imaginary Part } c_\alpha) \omega_\alpha a_{j\alpha}] \end{aligned} \tag{8-43}$$

Since the constants $a_{j\alpha}$ are all real, this pair of equations determines both the real and imaginary parts of c_α.

We can solve Eqs. (8-43) in a simple way, making use of the important property expressed by Eq. (8-48), which we now prove. For any particular α the constants $a_{j\alpha}$ satisfy

$$\omega_\alpha{}^2 a_{j\alpha} = \sum_{k=1}^{n} v_{jk} a_{k\alpha} \tag{8-44}$$

If we multiply this equation by $a_{j\beta}$ and sum over all values of j, we find

$$\omega_\alpha{}^2 \sum_{j=1}^{n} a_{j\alpha} a_{j\beta} = \sum_{k=1}^{n} \sum_{j=1}^{n} v_{jk} a_{k\alpha} a_{j\beta} \tag{8-45}$$

Since the coefficients v_{jk} are symmetrical, the left-hand side of Eq.

(8-45) will be the same if α and β are interchanged. This means

$$(\omega_\alpha^2 - \omega_\beta^2) \sum_{j=1}^{n} a_{j\alpha} a_{j\beta} = 0 \tag{8-46}$$

Thus if the frequencies ω_α and ω_β are different, it must be that

$$\sum_{j=1}^{n} a_{j\alpha} a_{j\beta} = 0 \tag{8-47}$$

If the two frequencies are the same, then the constants $a_{j\alpha}$ are not determinant. Instead, we have the freedom to make an arbitrary choice which can be made in such a way that Eq. (8-47) is satisfied for $\alpha \neq \beta$. Thus finally, making use of the normalization defined in Eq. (8-40), we can write

$$\sum_{j=1}^{n} a_{j\alpha} a_{j\beta} = \delta_{\alpha\beta} \tag{8-48}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta.

We can now easily find the real part of c_α from Eqs. (8-43). Multiply the first of Eqs. (8-43) by $a_{j\beta}$ and sum over values of α. All terms on the right-hand side vanish except that for $\alpha = \beta$, which gives

$$\text{R.P. } c_\beta = \sum_{j=1}^{n} a_{j\beta} q_j(0) \tag{8-49}$$

In a similar manner we can find

$$\text{I.P. } c_\beta = -\frac{1}{\omega_\beta} \sum_{j=1}^{n} a_{j\beta} \dot{q}_j(0) \tag{8-50}$$

Thus a complete description of any arbitrary motion of the system can be determined from a knowledge of the normal modes of the system and the initial conditions of the motion.

8-3 NORMAL COORDINATES

We can analyze the motion of the system in another way. Let us choose a new set of coordinates $Q_\alpha(t)$, which are a particular linear combination of the old coordinates, namely,

$$Q_\alpha(t) = \sum_{j=1}^{n} a_{j\alpha} q_j(t) \tag{8-51}$$

Alternatively, the old coordinates can be given in terms of the new by

$$q_j(t) = \sum_{\alpha=1}^{n} a_{j\alpha} Q_\alpha(t) \tag{8-52}$$

Making use of Eq. (8-48), we can write the kinetic energy of the system as

$$\text{K.E.} = \frac{1}{2} \sum_{j=1}^{n} \dot{q}_j^2 = \frac{1}{2} \sum_{j=1}^{n} a_{j\alpha} a_{j\beta} \dot{Q}_\alpha \dot{Q}_\beta = \frac{1}{2} \sum_{\alpha=1}^{n} \dot{Q}_\alpha^2 \tag{8-53}$$

The potential energy is

$$V = \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} v_{jk} q_j q_k = \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} v_{jk} a_{j\alpha} a_{k\beta} Q_\alpha Q_\beta \tag{8-54}$$

From Eq. (8-38) we have

$$\sum_{k=1}^{n} v_{jk} a_{k\beta} = \omega_\beta^2 a_{j\beta} \tag{8-55}$$

which means that the potential energy can be written as [using Eq. (8-48)]

$$V = \frac{1}{2} \omega_\beta^2 Q_\beta Q_\alpha = \sum_{j=1}^{n} a_{j\alpha} a_{j\beta} = \frac{1}{2} \sum_{\alpha=1}^{n} \omega_\alpha^2 Q_\alpha^2 \tag{8-56}$$

Thus the lagrangian of Eq. (8-34) can be written in terms of the new variables as

$$L = \frac{1}{2} \sum_{\alpha=1}^{n} (\dot{Q}_\alpha^2 - \omega_\alpha^2 Q_\alpha^2) \tag{8-57}$$

The lagrangian in this form represents a set of harmonic oscillators which no longer interact. That is, the variables are separated. Each oscillator has unit mass and its own particular frequency ω_α. The equation of motion for a particular oscillator is

$$\ddot{Q}_\alpha = -\omega_\alpha^2 Q_\alpha \tag{8-58}$$

This means that each mode oscillates freely at its own frequency independently of any other mode. By comparing Eqs. (8-49) and (8-50) with Eq. (8-51) we see that the real part of c_β and the imaginary part of $-c_\beta \omega_\beta$ are just the initial coordinate $Q_\beta(0)$ and the initial velocity $\dot{Q}_\beta(0)$, respectively, of the β mode. Thus the complicated molecule is equivalent to a simple set of independent harmonic oscillators.

This new set of coordinates Q_α which permit us to describe the system as a set of independent oscillators is called a set of *normal coordinates.* Using the lagrangian given by Eq. (8-57), we can write the path integral describing the motion of the system in terms of normal coordinates as

$$K = \int \left\{ \exp \left[\frac{i}{\hbar} \sum_{\alpha=1}^{n} \int (\dot{Q}_\alpha^2 - \omega^2 Q_\alpha^2)\, dt \right] \right\} \mathcal{D}Q_1\, \mathcal{D}Q_2 \cdots \mathcal{D}Q_n \tag{8-59}$$

This last result can be obtained directly from Eq. (8-35) by the explicit substitution $q_j(t) = \sum_\alpha a_{j\alpha} Q_\alpha(t)$. The exponent simplifies just as in the classical case, while $\mathcal{D}q_1 \cdots \mathcal{D}q_n = \mathcal{D}Q_1, \ldots, \mathcal{D}Q_n$ at least within a constant factor. [Since the transformation of coordinates is linear, the jacobian is constant. Any such constant can be contained in the definition of the normalization factors for the path integral $\mathcal{D}Q_1(t)\, \mathcal{D}Q_2(t) \cdots$.]

This form of the path integral can be broken down into a product of path integrals. Thus

$$K = \prod_{\alpha=1}^{n} \int \left\{ \exp \left[\frac{i}{2\hbar} \int (\dot{Q}_\alpha^2 - \omega^2 Q_\alpha^2)\, dt \right] \right\} \mathcal{D}Q_\alpha \tag{8-60}$$

where each path integral now describes only one mode and each mode is a simple one-dimensional oscillator, for which we have already obtained a solution. In this manner any problem of interacting harmonic oscillators can be analyzed.

Since the path integral for the kernel can be separated into a product of path integrals, it follows that the wave function for the system in a given energy state can be written as the product of wave functions of each mode as discussed in Sec. 3-8.

As shown in Sec. 8-1, the wave functions for each separate mode are proportional to $\exp(iE_n t/\hbar)$, where E_n is the energy of the mode. A product of such wave functions is then proportional to $\exp[(it/\hbar) \sum_n E_n]$. From this it follows that the total energy of the system of oscillators is equal to the sum of all the separate energies. The energy in the α mode is $\hbar\omega_\alpha(m_\alpha + \frac{1}{2})$, where m_α is an integer. The energy of the whole system is then

$$E = \hbar\omega_1(m_1 + \tfrac{1}{2}) + \hbar\omega_2(m_2 + \tfrac{1}{2}) + \cdots + \hbar\omega_n(m_n + \tfrac{1}{2}) \tag{8-61}$$

where $m_1, m_2, \ldots$ are all integers (including zero). All independent choices are allowable because the excitation of oscillator 1 and oscillator 2 can be in different degrees.

If $\phi_n(Q)$ is the harmonic oscillator wave function for the nth level [as given in Eq. (8-7)], then the wave function for the complete system is

$$\phi_{m_1}(Q_1)\phi_{m_2}(Q_2) \cdots \phi_{m_n}(Q_n) = \prod_{\alpha=1}^{n} \phi_{m\alpha}(Q_\alpha) \tag{8-62}$$

Each $\phi_{m\alpha}(Q_\alpha)$ is as given in Eq. (8-7) with ω replaced by ω_α. In this way classical physics, whereby we determine the normal modes, and quantum mechanics, whereby we determine the wave function and energy levels of a simple harmonic oscillator, are combined to give a complete solution for the energy levels and eigenfunctions of a polyatomic molecule.

We can express the wave functions in terms of the original coordinates $q_i(t)$ by using the transformation equation (8-51). For example, the lowest energy state of a system, which has the energy $\frac{\hbar}{2}\sum_{\alpha=1}^{n}\omega_\alpha$, has the wave function

$$\Phi_0 = \prod_{\alpha=1}^{n} \exp\left(-\frac{Q_\alpha{}^2}{2\omega_\alpha}\right) = \exp\left(-\frac{1}{2}\sum_{\alpha=1}^{n}\frac{Q_\alpha{}^2}{\omega_\alpha}\right)$$

$$= \exp\left(-\frac{1}{2}\sum_{\alpha=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{a_{j\alpha}a_{k\alpha}q_jq_k}{\omega_\alpha}\right) \tag{8-63}$$

That is, the wave function is an exponential function of the quadratic form $-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} M_{jk}q_jq_k$, where the matrix element M_{jk} is

$$M_{jk} = \sum_{\alpha=1}^{n}\frac{a_{j\alpha}a_{k\alpha}}{\omega_\alpha} \tag{8-64}$$

Problem 8-2 Show that the M_{jk} matrix is the reciprocal square root of the v_{jk} matrix. That is, show

$$\sum_{k=1}^{n}\sum_{l=1}^{n} M_{jk}M_{kl}v_{lm} = \delta_{jm} \tag{8-65}$$

It may happen that some of the frequencies ω are zero. For example, for the molecule CO_2 the modes 5 to 9, as pictured in Fig. 8-1, all have zero frequency. They correspond to a translation or a rotation of the whole molecule, motions for which there is no restoring force. Since there is no restoring force, the assumption that the coordinates

Q_α are small is not generally true. A more exact analysis of the translation or rotation kinetic energy must be undertaken. Since such motions are not of interest in the present discussion, we shall assume that these modes, and their coordinates, either do not exist or are never excited, so that we have been dealing with modes for which $\omega \neq 0$. If for particular values of α the solutions ω_α^2 come out negative (so that ω is imaginary), the system is in unstable equilibrium for motions in this mode, like a pencil balanced on its point. Instead of being simple harmonic, the motion is exponentially divergent, and again the coordinates Q_α do not stay small. This case again is of no interest in the present discussion, and we shall assume that there are no such modes.

8-4 THE ONE - DIMENSIONAL CRYSTAL

A Simple Model. We can think of a crystal as a large polyatomic molecule spread out in a three-dimensional array. We can begin learning about it by first studying a simpler one-dimensional line of equal atoms equally spaced, as in Fig. 8-2. Let the mass of each atom be 1 and let the displacement of the jth atom from its equilibrium position be q_j. We suppose that the motions are restricted to lie along the line of the array, i.e., longitudinal motions only. Next, suppose each atom interacts only with its neighbor and with a potential $V(R)$ which depends upon the separation R between them. That is, we suppose the atoms are connected together by a set of springs. The equilibrium separation gives a minimum value to the potential. We shall assign this minimum the value 0. Suppose ΔR is the difference between the equilibrium displacement and some particular displacement. We can expand the potential in a power series in terms of ΔR, in a manner analogous to that of Eq. (8-32). We shall restrict our attention to those displacements which are so small that all terms higher than the second order in this expansion can be neglected. Between the jth and $(j + 1)$st atoms the change in separation away from the equilibrium separations is $q_{j+1} - q_j = \Delta R_{j,j+1}$. We shall

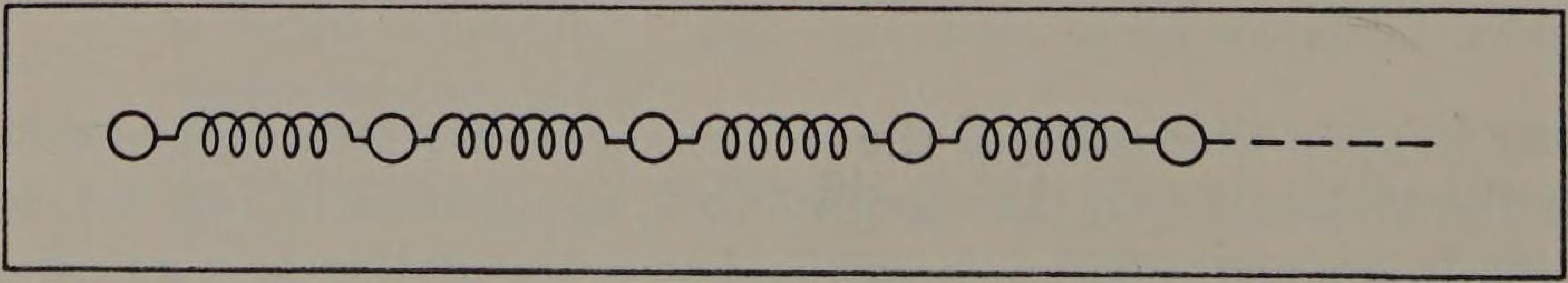


Fig. 8-2 A model of a one-dimensional "crystal," with mass particles evenly spaced along a line and springs connecting neighboring particles.

call the second derivative of the potential with respect to the displacement ν^2 (the same for all atoms in the string). Then the potential energy associated with such a displacement is

$$V_{j,j+1} = \tfrac{1}{2}\nu^2(q_{j+1} - q_j)^2 \tag{8-66}$$

and the lagrangian can be written as

$$L = \sum_{j=1}^{N} \frac{1}{2}\dot{q}_j^2 - \sum_{j=1}^{N-1} \frac{\nu^2}{2}(q_{j+1} - q_j)^2 \tag{8-67}$$

If the first and last atoms are unattached, then the term for $j = N$ in the expression for potential energy must be omitted.

Based on this lagrangian the equations of motion for the atoms along the line are

$$\ddot{q}_j = \nu^2[(q_{j+1} - q_j) - (q_j - q_{j-1})] \tag{8-68}$$

for all j except the end points $j = 1$ and $n = N$. Now this fact that the end particles have to be given separate consideration is just a minor annoyance for most problems. Usually we are interested in the gross properties for a large solid and are not concerned with surface or boundary effects. In such cases the main results desired are really independent of the actual boundary conditions (e.g., whether or not the end atoms are left free or are tied down, etc.). To avoid this problem, theoretical physicists use a trick of assuming a special set of simple boundary conditions, called periodic boundary conditions, so that these end points do not require special consideration in the analysis. Unfortunately, these special boundary conditions occur in actuality only rarely, if at all, but for phenomena which are independent of boundary effects the trick is useful.

The idea is to imagine that the string of atoms goes on beyond N, but a displacement of the $(N + j)$th atom is always exactly equal to that of the jth atom. Thus the boundary condition is

$$q_{N+1} = q_1 \qquad \dot{q}_{N+1} = \dot{q}_1 \tag{8-69}$$

This boundary condition would be right if our string were tied in a circle like a pearl necklace. However, in three dimensions there is no such picture to represent the boundary condition, and it must be considered completely artificial.

The value of this particular boundary condition is this. Most general ways of terminating the string (e.g., tying the last atom to a rigid wall, leaving the last atom free, etc.) result in a reflection of any wave traveling down the string. Only if the last atom is tied to another string of atoms of identical characteristics will no such reflection occur. Thus the boundary condition is analogous to tying a transmission line to a characteristic impedance in order to avoid reflections. The characteristic impedance is equivalent to an infinity of more line. In the present case we accomplish this end by tying the string to itself. We call the boundary conditions periodic, because anything that happens at the point k in the string is repeated again at the point $N + k$ and again at $2N + k$, etc. With this boundary condition Eq. (8-68) for the motion of the atoms is valid for all of the atoms.

Solving the Classical Equations of Motion. We assume that the displacements q are periodic with frequency ω. Then we must solve

$$\omega^2 q_j = \nu^2(q_{j+1} - 2q_j + q_{j-1}) \tag{8-70}$$

We could write down this set of equations in a determinant, and it turns out that the determinantal equation so obtained can be evaluated by theorems in mathematics. But this just means that the equations can be solved directly, and it is easier to solve them that way.

We shall restrict the symbol i to mean $\sqrt{-1}$, and not use it for an index. Solutions are of the form

$$q_j = Ae^{i(j\beta - \omega t)} = a_j e^{-i\omega t} \tag{8-71}$$

where β is a constant taking on a discrete set of values. This solution can be verified by substituting into Eq. (8-70). The frequency is given by

$$\omega^2 = \nu^2(e^{i\beta} - 2 + e^{-i\beta}) = 4\nu^2 \sin^2\frac{\beta}{2} \tag{8-72}$$

This gives the values of ω in terms of β, but not all values of β are allowed. The periodic boundary condition implies that $\beta = 2\pi\alpha/N$, where α is an integer $= 0, 1, 2, \ldots, N-1$. (The case $\alpha = 0$ is just a translation, and we can omit it if desired. Furthermore, the case given by $\alpha = N + \alpha'$ is the same as the case given by α'.) Thus for any particular choice of α we have the frequency

$$\omega_\alpha = 2\nu \sin\frac{\pi\alpha}{N} \tag{8-73}$$

and the amplitude for the jth coordinate at that frequency is

$$a_{j\alpha} = Ae^{i2\pi\alpha j/N} \tag{8-74}$$

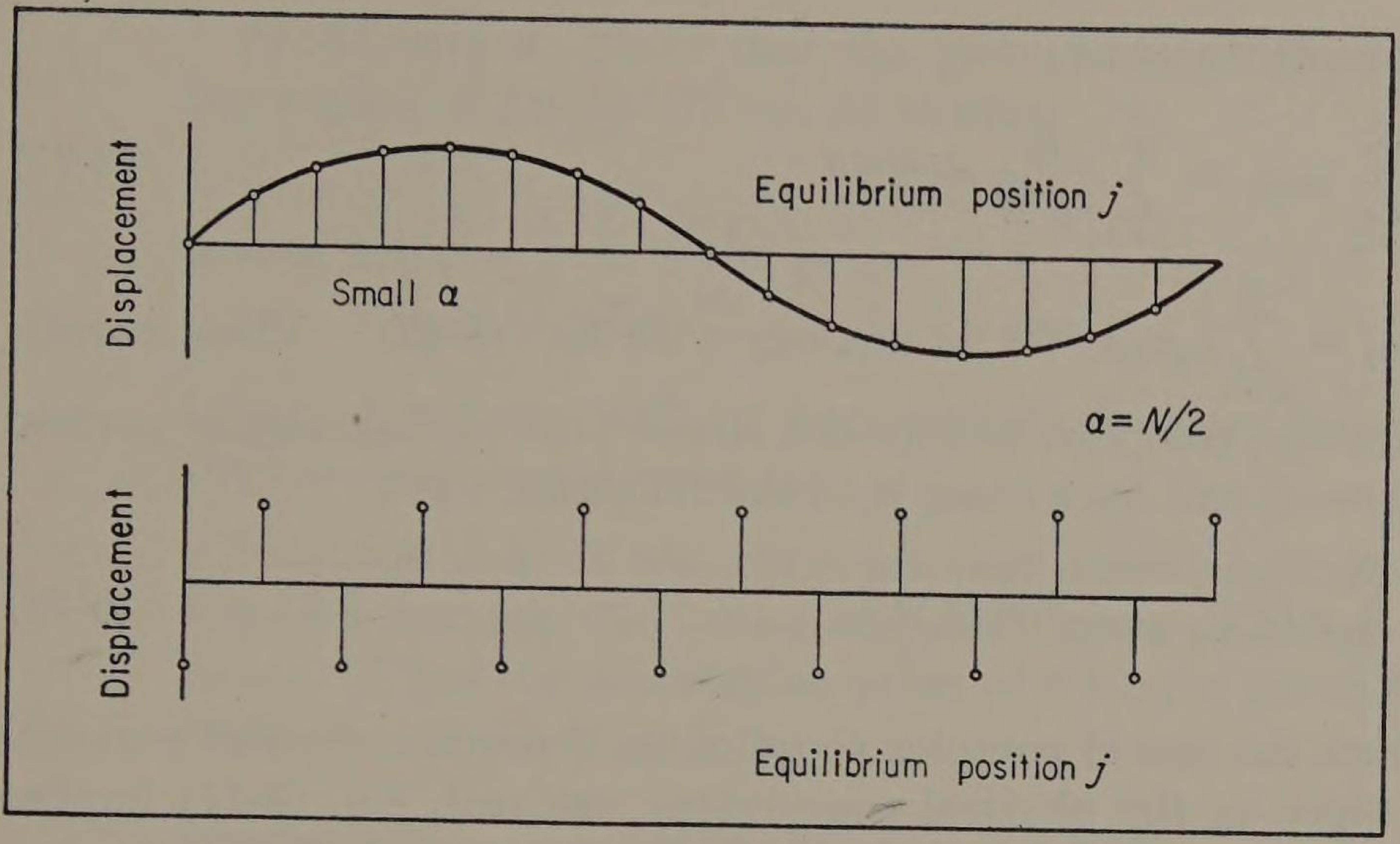


Fig. 8-3 The displacement of atoms along a string is plotted as the ordinate against the equilibrium positions j equally spaced along the abscissa. In the upper case the wavelength is long compared to the spacing between atoms (small α). In the lower case $\alpha = N/2$, and the displacements no longer give the appearance of a smooth sine wave.

The constants $a_{j\alpha}$ determined in the last equation are complex. They could be made real by combining solutions for α and $-\alpha$ (or α and $N - \alpha$). However, it is more convenient to leave them in the complex form. It is convenient to consider both positive and negative values of α, however, and so if N is odd, for example, to consider the range of α to be $-\frac{1}{2}(N - 1)$ to $+\frac{1}{2}(N - 1)$ rather than 0 to $N - 1$.

The relative displacements of the atoms in the string depend on the size of α. The situation for two values of α, one with α small and the other with $\alpha = N/2$, is shown in Fig. 8-3.

Although the comparative magnitudes of the various constants $a_{j\alpha}$ are determined by Eq. (8-74), the general magnitude, determined by the constants A, is still arbitrary. We establish this with a normalizing equation analogous to Eq. (8-48). Thus choose A so that

$$\sum_{j=1}^{N} a_{j\alpha}^{*} a_{j\beta} = \delta_{\alpha\beta} \tag{8-75}$$

which implies that

$$A = \frac{1}{\sqrt{N}} \tag{8-76}$$

We are now in a position to represent the various modes with their normal coordinates as

$$Q_\alpha = \sum_{j=1}^{N} a_{j\alpha} q_j = \sum_{j=1}^{N} \frac{q_j}{\sqrt{N}} e^{ij\cdot 2\pi\alpha/N} \tag{8-77}$$

where $q_j = \sum_{\alpha=1}^{N} c_\alpha a_{j\alpha} e^{-i\omega_\alpha t}$, in analogy with Eq. (8-42). These coordinates are also complex, but we can ensure that the lagrangian derived from them is real by writing it in the following way:

$$L = \frac{1}{2} \sum_{\alpha=1}^{N} (\dot{Q}^*_\alpha \dot{Q}_\alpha - \omega_\alpha{}^2 Q^*_\alpha Q_\alpha) \tag{8-78}$$

Perhaps this use of complex coordinates Q needs a word of explanation. Since q_j, the physical coordinates, are real, Eq. (8-77) implies $Q^*_\alpha = Q_{-\alpha}$ so that, although two real numbers are really required to specify each complex coordinate Q_α, only N independent real numbers are needed for all of them. If one prefers real coordinates, one can define instead two real quantities as coordinates by writing

$$Q_\alpha = (Q_\alpha{}^c - iQ_\alpha{}^s) \frac{1}{\sqrt{2}}$$

$$Q_\alpha{}^c = \frac{1}{\sqrt{2}} (Q_\alpha + Q_{-\alpha}) \tag{8-79}$$

$$Q_\alpha{}^s = \frac{i}{\sqrt{2}} (Q_\alpha - Q_{-\alpha}) \tag{8-80}$$

but the values of α now run the range only from 0 to $N - 1$. Then a term such as the kinetic energy is expressed in real variables as

$$\tfrac{1}{2}[(\dot{Q}_\alpha{}^c)^2 + (\dot{Q}_\alpha{}^s)^2] = \dot{Q}_\alpha \dot{Q}_{-\alpha} = \dot{Q}_\alpha \dot{Q}^*_\alpha \tag{8-81}$$

[The factor ½ reappears in Eq. (8-78) because we sum there over all α, plus and minus, counting thereby each term twice, $Q^*_{-\alpha} Q_{-\alpha} = Q_\alpha Q^*_\alpha$.] Thus quadratic expressions derived previously for real quantities appear now as products of one complex number by its conjugate [for example, Eq. (8-75)].

Problem 8-3 Show that $Q_\alpha{}^c$, $Q_\alpha{}^s$ are normal coordinates, representing, however, standing waves $\sqrt{2}\cos(2\pi\alpha j/N)$ and $\sqrt{2}\sin(2\pi\alpha j/N)$, respectively, i.e.,

$$q_j = \sum_{\alpha=0}^{\frac{1}{2}(N-1)} Q_\alpha{}^c \sqrt{2} \cos\frac{2\pi\alpha j}{N} + \sum_{\alpha=0}^{\frac{1}{2}(N-1)} Q_\alpha{}^s \sqrt{2} \sin\frac{2\pi\alpha j}{N} \tag{8-82}$$

for N odd.

Problem 8-4 Show that the ground-state wave function for the lagrangian of Eq. (8-78) can be written

$$\Phi = A \exp\left(-\frac{1}{2}\sum_{\alpha=0}^{N} \frac{Q_\alpha^* Q_\alpha}{\omega_\alpha}\right) \tag{8-83}$$

(where A is a constant) by starting with the wave function in terms of the real variables $Q_\alpha{}^c$ and $Q_\alpha{}^s$.

Problem 8-5 A transition element which employs the same wave function as both the initial and final states is called an *expectation value.*[1] Thus the expectation value of F for the ground state Φ of Eq. (8-83) is

$$\langle\Phi_0|F|\Phi_0\rangle = \int \cdot\cdot\cdot \int \Phi_0^* F \Phi_0 \, dQ_0 \, dQ_1 \cdot\cdot\cdot dQ_N \tag{8-84}$$

Show that the following expectation values are correct:

$$\begin{aligned}
\langle\Phi_0|Q_\alpha|\Phi_0\rangle &= \langle\Phi_0|Q_\alpha^*|\Phi_0\rangle = 0 \\
\langle\Phi_0|Q_\alpha{}^2|\Phi_0\rangle &= \langle\Phi_0|Q_\alpha^{*2}|\Phi_0\rangle = 0 \\
\langle\Phi_0|Q_\alpha^* Q_\alpha|\Phi_0\rangle &= \frac{1}{2\omega_\alpha}\langle\Phi_0|1|\Phi_0\rangle \\
\langle\Phi_0|Q_\alpha^* Q_\beta|\Phi_0\rangle &= 0 \text{ if } \alpha \neq \beta
\end{aligned} \tag{8-85}$$

Thus with the lagrangian written in terms of the normal coordinates we have reduced the system to a set of independent simple harmonic oscillators. The quantum-mechanical part of the solution follows in a straightforward manner just as it did for the case of the polyatomic molecule. All that we need to know is the quantum-mechanical solution for a simple independent harmonic oscillator.

Problem 8-6 Show that the constants $a_{j\alpha}$ are the same even if the coupling is not just to the nearest neighbors but extends with strengths λ_k to an atom k spaces away. Assuming λ_k falls rapidly enough for large k, find the values of the frequency ω_α when such a coupling is present, i.e., when the potential energy, instead of being given by Eq. (8-66), is given by a similar equation, but one which contains the relative displacements of all pairs of atoms, each one multiplied by the appropriate λ_k, that is, $V = (\nu^2/2) \sum_k \sum_j \lambda_k (q_{k+j} - q_j)^2$.

---

[1] Compare this definition of *expectation value* with the definition of the *expected value* of an operator given in Sec. 5-3, particularly in Eq. (5-46).

8-5 THE APPROXIMATION OF CONTINUITY

The particular modes which we have determined here are those in which each atom oscillates with a phase difference behind the one next in line. There is a wave of oscillation passing down the line of atoms. If the phase difference between adjacent atoms is small, then the wavelength is long.

Of special interest is the behavior of the atoms in the long-wavelength modes. If the wavelength greatly exceeds the spacing between the atoms, this spacing is unimportant. In this case the motion can be very well described by the fictitious "continuous medium" concept. A line of atoms can be replaced by a continuous rod with certain average properties, such as the mass per unit length $\rho = 1/d$ (recall that we have assigned unit mass to each atom). More physically, a real rod is actually a discrete set of atoms. In this section we shall develop the approximation of continuity, wherein a line of atoms is replaced by a continuous string.

For a particular mode of motion the phase difference between adjacent atoms is $2\pi\alpha/N$, so that a wavelength contains N/α atoms, or if d is the equilibrium separation distance between neighboring atoms, the wavelength is $\lambda = Nd/\alpha$. The wave number is

$$k = \frac{2\pi}{\lambda} = \frac{2\pi\alpha}{Nd} \tag{8-86}$$

The wave aspect is made more clear in the mathematical representation of the motion by a slight change of notation. We shall refer to each mode by its k value instead of by its α value. Then a summation α over the modes means a sum over discrete values of k. These values are the integers multiplied by $2\pi/L$, where $L = Nd$ is the length of the string. Suppose $x_j = jd$ is the equilibrium position of the jth atom. Then the equations describing the motion of the atom become

$$a_{jk} = \frac{1}{\sqrt{N}} e^{ikx_j} \tag{8-87}$$

$$Q_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} q_j e^{ikx_j} \tag{8-88}$$

$$q_j = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} Q_k e^{-ikx_j} \tag{8-89}$$

and

$$\omega_k = 2\nu \sin \frac{kd}{2} \tag{8-90}$$

We now assume that the separation between atoms is very small compared to the length over which disturbances change. Using the symbols we have already defined, such situations as this are described by $kd \ll 1$. If we call the product $\nu d = c$, then for kd small we have $\omega \simeq kc$. In this situation we can think of the coordinates q_j as being functions of position along the line of atoms. That is, we can specify the displacement of the jth atom, as shown in Fig. 8-3. For long waves the displacements $q(x_j)$ and $q(x_j + 1)$ are nearly equal, and we can consider the function $q(x)$ as a smooth continuous function defining displacement as a function of position along the line. The normal coordinate Q_k is a Fourier transform of $q(x)$. That is, Eq. (8-88) can be replaced by

$$Q(k) = \frac{\sqrt{N}}{L} \int_0^L q(x) e^{ikx}\, dx \tag{8-91}$$

This replacement is based upon the approximate relation

$$\sum_{j=1}^{N} (\quad)_j \simeq \frac{N}{L} \int_0^L (\quad)\, dx \tag{8-92}$$

which becomes more valid as the spacing between discrete points becomes very small.

A similar relation, namely,

$$\sum_{k=1}^{N} (\quad)_k \simeq \frac{L}{2\pi} \int_0^{2\pi/d} (\quad)\, dk \tag{8-93}$$

leads to the inverse transform

$$q(x) = \frac{L}{2\pi\sqrt{N}} \int_0^{2\pi/d} Q(k) e^{-ikx}\, dk \tag{8-94}$$

To make these quantities of more direct physical significance, let the actual displacement of the jth atom be u_j. That is, $q_j = \sqrt{m}\, u_j$, where m is the mass of one atom and is equal to ρd. Let the Fourier transform of u be U. Thus

$$U(k) = \int_0^L u(x) e^{ikx}\, dx \tag{8-95}$$

while the inverse transform is

$$u(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} U(k)e^{-ikx}\,dk \tag{8-96}$$

The new normal coordinate is then $U(k)$, and it is related to the previous normal coordinate $Q(k)$ by

$$U(k) = \frac{\sqrt{mL}}{\sqrt{N}}\,Q(k) \tag{8-97}$$

The expression for the kinetic energy in terms of $u(x,t)$ can be worked out with the help of Eq. (8-92) to be

$$\text{K.E.} = \frac{1}{2}\int \rho\left(\frac{\partial u}{\partial t}\right)^2 dx \tag{8-98}$$

To determine the potential energy in terms of all the new variables, we need to express the difference in the displacements of two adjacent atoms as a continuous function of position. Using our approximation of continuity, we can write

$$q_{j+1} - q_j = \sqrt{m}\,[u(x_{j+1},t) - u(x_j,t)] \simeq d\sqrt{m}\,\frac{\partial u}{\partial x} \tag{8-99}$$

That means that the potential energy is

$$V = \frac{\nu^2 d^2}{2}\frac{m}{d}\int_0^L \left(\frac{\partial u}{\partial x}\right)^2 dx = \frac{\rho c^2}{2}\int_0^L \left(\frac{\partial u}{\partial x}\right)^2 dx \tag{8-100}$$

In the last equation we have used the constant $c = \nu d$. This constant is actually a measure of the elasticity. We can define it physically in the following manner: Suppose we stretch the line of atoms, which has length L, by a fractional increase of amount ϵ, that is, to the new length $L(1 + \epsilon)$. (We are considering a static stretch, not a vibration.) This means that we make the separation between each pair of atoms equal to $d(1 + \epsilon)$ instead of d. Thus the difference in displacements of adjacent atoms becomes

$$q_{j+1} - q_j = \epsilon d\sqrt{m} \tag{8-101}$$

Using Eq. (8-66), this means potential energy put into the string by the stretching is

$$V = \frac{\nu^2}{2}\epsilon^2 d^2 mN = \frac{\rho c^2}{2}\epsilon^2 L \tag{8-102}$$

Thus the force that is needed to stretch the string is, in the limit of small ϵ,

$$\frac{V}{\epsilon L} = \rho c^2 \epsilon \tag{8-103}$$

This last equation gives the stress in the string, and the strain (stretch per unit length) is, of course, ϵ. Thus we have

$$\frac{\text{Stress}}{\text{Strain}} = \rho c^2 = \text{elastic constant} \tag{8-104}$$

Combining Eqs. (8-98) and (8-100), we can construct the lagrangian as

$$L = \frac{\rho}{2} \int \left(\frac{\partial u}{\partial t}\right)^2 dx - \frac{\rho c^2}{2} \int \left(\frac{\partial u}{\partial x}\right)^2 dx \tag{8-105}$$

The fundamental modes which we are considering have the form e^{ikx}, and the normal coordinates are $U(k,t)$. The reader can show that the lagrangian can be expressed in terms of these normal coordinates as

$$L = \frac{\rho}{2} \int \left[\frac{\partial U(k,t)}{\partial t}\right]^2 \frac{dk}{2\pi} - \frac{\rho c^2}{2} \int k^2 U^2(k,t) \frac{dk}{2\pi} \tag{8-106}$$

We can consider the system described by this lagrangian as a set of harmonic oscillators, one oscillator for each value of k. In our present approximation of continuity, k is a continuous variable with an infinite number of values. We can reintroduce the picture of discrete atoms by remembering that the integral over dk is really a sum over discrete values of k, where the various discrete values of k are spaced a distance $2\pi/L$ apart, with L the length of the string, and the number of such values is equal to the number of atoms in the string.

We can get equations of motion in terms of the continuous variables by finding the extremum of the action integral $\int_0^T L\, dt$. Using the form of L given by Eq. (8-105), the resulting equation of motion is

$$\rho \frac{\partial^2 u}{\partial t^2} = \rho c^2 \frac{\partial^2 u}{\partial x^2} \tag{8-107}$$

Following a line of argument demonstrated by Eq. (8-99), we can see that this equation of motion is analogous to the previous equation of motion which we derived, namely, Eq. (8-68). Equation (8-107) has the solution

$$u = e^{-i\omega t} a(x) \tag{8-108}$$

in analogy with Eq. (8-71), where

$$-\omega^2 a = c^2 \left(\frac{da}{dx}\right)^2 \tag{8-109}$$

in analogy with Eq. (8-70), and

$$a(x) = e^{ikx} \tag{8-110}$$

in analogy with Eq. (8-74).

Combining Eqs. (8-109) and (8-110), we see that $\omega = kc$. This is the analogue of Eq. (8-90), and, as a matter of fact, in the limit of small k, Eq. (8-90) reduces to this relation.

The motion described by Eq. (8-108), with the value of a given by Eq. (8-110), is that of the traveling wave moving with velocity c. That is to say, c is the speed of sound along the line of atoms. Actually, a real system shows dispersion; that is, ω is not proportional to k. For wavelengths which are of the same order as the atomic spacing this lack of proportionality becomes important, as shown by Eq. (8-90).

8-6 QUANTUM MECHANICS OF A LINE OF ATOMS

The behavior of the atoms in a string can be described in terms of modes of motion. Each mode is a harmonic oscillator. The energy state of any particular mode is determined by the quantum number for that mode. Each mode is identified by its wave number k or its frequency ω. A mode of frequency ω can have the energy values $\hbar\omega/2$, $3\hbar\omega/2$, $5\hbar\omega/2$, . . . , or 0, $\hbar\omega$, $2\hbar\omega$, . . . , above the ground state energy of $\hbar\omega/2$. For these cases we would say that there are 0, 1, 2, . . . *phonons* of wave number k (or frequency ω) present.

It is possible to have several different modes excited simultaneously. For example, we could have (1) the mode of wave number k_1 excited to its first level above the ground state, (2) the mode of wave number k_2 excited to its first level also, and (3) the mode of wave number k_3 excited to its second state above the ground level. The state of the complete system would then have the total energy $h(\omega_1 + \omega_2 + 2\omega_3)$ above the ground state. We would say that there are four phonons present, one phonon of wave number k_1, one of wave number k_2, and two of wave number k_3.

The ground state of the entire system has the energy

$$E_{gnd} = \sum_k \frac{\hbar\omega_k}{2} \tag{8-111}$$

Using the approximation of continuity and letting $\omega = kc$, this becomes

$$E_{gnd} = \frac{L}{2\pi} \int_0^{k_{\max}} \frac{\hbar kc}{2}\, dk \tag{8-112}$$

We notice that if the upper limit $k_{\max}$ on the integral over k goes to infinity, then the integral diverges. However, the form $\omega = kc$ which we have used in this expression is valid only for long waves (i.e., small values of k).

We can make a better determination of the ground-state energy by using the correct expression for ω and establishing a reasonable upper limit for the integral over k. Thus, using Eq. (8-90) for ω_k, we can write the ground-state energy as

$$E_{gnd} = \sum_{k=-k_{\max}}^{k=k_{\max}} \frac{\hbar}{2} \nu \sin \frac{kd}{2} \tag{8-113}$$

where

$$k_{\max} = \frac{2\pi}{d} \tag{8-114}$$

This can be rewritten as

$$\sum_{n=-N/2}^{n=N/2} \hbar\nu \left| \sin \frac{\pi n}{N} \right| = 2\hbar\nu(\text{I.P.}) \sum_{n=0}^{N/2} e^{i\pi n/N} \tag{8-115}$$

For a very large N this result can be approximated by

$$E_{gnd} = 2\hbar c L \frac{1}{\pi d^2} \tag{8-116}$$

The results show that the energy is proportional to the length of the string, but apparently the energy has no limit as the spacing d approaches zero. That is, the ground-state energy is infinite for a continuous medium. Of course, for real matter the energy is finite.

It is very convenient to measure, instead of the total energy, the excess energy above the ground state. There are two reasons for this: (1) Really, the ground-state energy is not known, nor is it usually interesting to the physical problem in question. For example, the true ground-state energy includes all of the energy of the electrons attached to the atoms. (2) When dealing with the excitation of only long waves, the approximation of continuity is very useful, and it gives a good approximation to the excitation energies. However, this approximation gives an invalid result for the ground-state energy, since it neglects the separation d (i.e., treats d as 0). Thus, we must avoid the necessity of evaluating the ground-state energy if we are to use the approximation of continuity.

8-7 THE THREE - DIMENSIONAL CRYSTAL

There is no difference in principle between a realistic three-dimensional crystal and the one-dimensional example which we have been considering. However, the detailed evaluation of the various modal frequencies is much harder. Results can be obtained in terms of the wave number $\mathbf{k}$, which is now a vector with the components k_x, k_y, and k_z. The frequency, written in terms of these components, is generally very complicated. There is more than one solution for each value of $\mathbf{k}$ because of the possibility of various polarizations (directions of vibration). Furthermore, a real crystal often consists not of an array of atoms equally spaced, but rather of an array of unit cells, each unit cell consisting of a group of atoms in some characteristic geometrical arrangement. If there are several atoms (say p) in such a unit cell (and this example can be illustrated with a one-dimensional approximation), then there are $3p$ frequencies for each value of $\mathbf{k}$.

In the three-dimensional crystal we can still use the approximation of continuity to good advantage. In this approximation the true lattice structure of the crystal generally makes itself felt through the existence of different properties in different directions (e.g., anisotropic compressibility). The symmetry of the lattice is reflected by the symmetry of elastic constants. Furthermore, the fundamental modes have vibration directions (polarization directions) which are not necessarily either parallel to or perpendicular to the direction of propagation of the wave.

For the present discussion, we shall assume that our substance displays the same elastic constants in all directions. (In general, it is not necessary for any crystal, even one as symmetric as a cubic crystal, to do this.) Then the waves are of two kinds, longitudinal and transverse. These two kinds of waves have different wave velocities, which we shall label c_L for the longitudinal and c_T for the transverse. For each $\mathbf{k}$ there are three modes. One of these has the frequency $\omega_L = c_L k$ (where k is the absolute magnitude of $\mathbf{k}$). Since, by hypothesis, there is no directional effect, the frequency is a function only of the absolute magnitude of the wave number and does not depend upon its specific components. There are two transverse modes (i.e., modes in which the direction of motion of the atoms is perpendicular to the direction of motion of the wave), both of which have the frequency $\omega_T = c_T k$.

Every separate mode, and that includes every separate direction of polarization, behaves like an independent oscillator.

Suppose we are dealing with the crystal of volume V. Let us compute the number of modes whose wave numbers lie in the k-space volume element $d^3\mathbf{k} = dk_x\, dk_y\, dk_z$ centered about the point $\mathbf{k}$. We assume the crystal is rectangular with edge lengths L_x, L_y, and L_z. We use the results obtained from the one-dimensional example to see that the discrete values of k_x are spaced apart a distance $2\pi/L_x$. So, in the range of length dk_x there are $dk_x L_x/2\pi$ discrete values of k_x. Applying this same reasoning to the other directions, we find that the number of discrete values of $\mathbf{k}$ included in the interval is

$$\frac{dk_x\, dk_y\, dk_z}{(2\pi)^3} L_x L_y L_z = \frac{d^3\mathbf{k}}{(2\pi)^3} V \tag{8-117}$$

This same result is obtained (in the limit of large crystals) for any shape.

For the general case the modal frequency $\omega_\mathbf{k}$ is, as we have mentioned, a very complicated function of $\mathbf{k}$ with several branches (values for the same $\mathbf{k}$), but its determination is a problem of classical physics; then the forms of oscillation in the fundamental modes are known, as are the normal coordinates describing these modes. The quantum-mechanical problem is then reduced to the solution of a simple set of oscillators, and all the properties can be worked out easily. The excitation of each mode is called the excitation of a phonon.

As a very simple special example, we shall consider the longitudinal modes of oscillation in an isotropic solid (i.e., sound or, in particular, longitudinal sound). We can start as we did in the one-dimensional example with the atoms in the crystal discretely spaced and later pass to the long-wavelength limit, or approximation of continuity.

A complete solution would show us all the effects of dispersion, the complex branches, and the transverse waves. It is a very interesting study. However, one need not carry out all of the steps in order to obtain the proper quantum-mechanical form of the continuity approximation. One can make use directly of the results of classical physics. The entire procedure, starting with discretely spaced point masses, then passing to the long-wavelength limit, is just as useful and just as valid in quantum mechanics as it is in classical physics. The lagrangian has the same form so long as one restriction is imposed, i.e., that the potential can be adequately represented by a quadratic function of the displacement. The reason for the similarity between the results of the classical and quantum-mechanical approach is that the procedure consists only of various linear transformations, e.g., transforming to normal coordinates followed by certain approximations, such as the approximation of continuity. These transformations and approximations can be done in quantum mechanics just exactly as they are done in classical physics.

The equations derived from classical physics are as follows. Suppose $u(\mathbf{r},t)$ represents the displacement of a particle whose equilibrium position is at $\mathbf{r}$. We assume that we are working in the long-wavelength region, so that the approximation of continuity applies. A plane-wave mode is easiest to describe in terms of the Fourier transform given by

$$\mathbf{U}(\mathbf{k},t) = \int_V \mathbf{u}(\mathbf{r},t)e^{i\mathbf{k}\cdot\mathbf{r}}\,d^3\mathbf{r} \tag{8-118}$$

where $\mathbf{r}$ is a spatial vector having the components x, y, z. The normal coordinates of the various modes depend on the relationship between the direction of $\mathbf{U}$ and the direction of the vector $\mathbf{k}$. That is, the coordinate $U_x(\mathbf{k},t)$ of the vector $\mathbf{U}$ does not necessarily represent a normal mode. For an isotropic material the three modes of a given $\mathbf{k}$ have the following normal coordinates:

$$U_1(\mathbf{k},t) = \frac{\mathbf{k}\cdot\mathbf{U}}{k} \tag{8-119}$$

(that is, the component of U in the direction of k) and

$$U_2(\mathbf{k},t) = \mathbf{e}_1\cdot\mathbf{U} \tag{8-120}$$

$$U_3(\mathbf{k},t) = \mathbf{e}_2\cdot\mathbf{U} \tag{8-121}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are two unit vectors perpendicular to $\mathbf{k}$ and perpendicular to each other. For the present study we shall restrict our attention to just that part of the kinetic and potential energy which arises from the longitudinal modes given by Eq. (8-119) and omit the transverse oscillations.

Using the results of classical physics the lagrangian for the longitudinal modes can be written as

$$L = \frac{\rho}{2}\iiint\left\{\left[\frac{\partial U_1(\mathbf{k},t)}{\partial t}\right]^2 - c^2k^2[U_1(\mathbf{k},t)]^2\right\}\frac{d^3\mathbf{k}}{(2\pi)^3} \tag{8-122}$$

Here we have introduced the speed of sound $c = \omega/k$, which is a function of the direction of propagation. This is a direct generalization from the one-dimensional example. In terms of the original variables

$\mathbf{u}(\mathbf{r},t)$ the lagrangian is

$$L = \frac{\rho}{2}\iiint\left[\left|\frac{\partial \mathbf{u}}{\partial t}\right|^2 - c^2\,(\nabla\cdot\mathbf{u})^2\right]d^3\mathbf{r} \tag{8-123}$$

The first term on the left-hand side of this equation is the kinetic energy, given by one-half the mass times the square of the velocity. The second term is the energy of compression given by $\nabla\cdot\mathbf{u}$, which is the compressional strain. No energy of sheer strain is included here because we have disregarded transverse elastic waves.

Variation of the lagrangian with respect to $\mathbf{u}$ produces the classical equations of motion as

$$\frac{1}{c^2}\frac{\partial^2\mathbf{u}}{\partial t^2} = -\nabla(\nabla\cdot\mathbf{u}) \tag{8-124}$$

If we define a compressional strain function equal to the divergence of $\mathbf{u}$, that is,

$$\phi = \nabla\cdot\mathbf{u} \tag{8-125}$$

we have the result

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = -\nabla^2\phi \tag{8-126}$$

which is the classical wave equation.

The Fourier transform of Eq. (8-124), using the kernel $\exp(i\mathbf{k}\cdot\mathbf{r})$ and taking the component of the result parallel to $\mathbf{k}$, gives

$$-\frac{1}{c^2}\frac{\partial^2 U_1}{\partial t^2} = k^2 U_1 \tag{8-127}$$

This is the equation of a single harmonic oscillator, and it shows us that $U_1(\mathbf{k},t)$ is indeed a normal coordinate.

The quantum-mechanical results from the lagrangian given by Eq. (8-123) can be obtained easily. The energy levels of the mode in question are given by $n\hbar(kc)$ above the ground state. Let us ask for the amplitude to go from a fixed initial set of coordinates $\mathbf{u}(\mathbf{r},0)$ to a given final set of coordinates $\mathbf{u}(\mathbf{r},T)$. It is

$$K[\mathbf{u}(\mathbf{r},T),T;\mathbf{u}(\mathbf{r},0),0] = \int\left(\exp\left\{-\frac{i\rho}{2\hbar}\int_0^T\int_0^{L_z}\int_0^{L_y}\int_0^{L_x}\left[\left|\frac{\partial\mathbf{u}}{\partial t}\right|^2 - c^2(\nabla\cdot\mathbf{u})^2\right]d^3\mathbf{r}\,dt\right\}\right)\mathcal{D}^3\mathbf{u}(\mathbf{r},t) \tag{8-128}$$

The path integral of Eq. (8-128) is carried out over the paths $\mathbf{u}(\mathbf{x},t)$ defined in terms of all three components of the vector $\mathbf{r}$, as well as the time t. It is subject, of course, to the condition that the function $\mathbf{u}(\mathbf{r},t)$ take on a given form at both the initial and final points. This is an interesting extension of our original path integral idea. Up to now we have dealt with integrands which were functionals of one (or perhaps a few) function $x(t)$ of one variable t, and we have carried out the integration over all such paths, or functions. Now we must integrate a functional of the function $\mathbf{u}(\mathbf{r},t)$ of four variables x, y, z, and t and carry out the path integration over all values of this function. We can accomplish this by the regular techniques which we have described before, for our integrand is still a gaussian functional.

The first step in the solution of the path integral is to find the path which leads to a stationary value for the integral appearing in the exponent, the one which satisfies Eq. (8-123), or, more conveniently, the wave equation given by Eq. (8-126). We must impose the required boundary conditions at the times $t = 0$ and $t = T$. Satisfying the boundary conditions is not a difficult problem; however, it is a little different from the usual problem in classical physics in which the coordinate and its derivative are given at $t = 0$, that is, $\mathbf{u}(\mathbf{r},0)$ and $(\partial\mathbf{u}/\partial t)_{t=0}$.

We could proceed along this line and solve the problem. However, we have learned from previous examples that it is much easier to transform the problem into normal coordinates before carrying out the path integral. Such a transformation gives us

$$K = \int_{U_1(0)}^{U_1(T)} \left\{\exp\left[-\frac{i\rho}{2\hbar}\sum_{\mathbf{k}}\int(\dot{U}_1^2 - k^2c^2U_1^2)\,dt\right]\right\}\mathfrak{D}U_1(\mathbf{k},t) \tag{8-129}$$

where the boundary values are given by

$$\begin{aligned} U_1(T) &= U_1(\mathbf{k},T) = \frac{\mathbf{k}}{k}\cdot\iiint e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}(\mathbf{r},T)\,d^3\mathbf{r} \\ U_1(0) &= U_1(\mathbf{k},0) = \frac{\mathbf{k}}{k}\cdot\iiint e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}(\mathbf{r},0)\,d^3\mathbf{r} \end{aligned} \tag{8-130}$$

This is once more the simpler type of path integral, where the path is described in terms of only the one variable t. Since the path integral can be written as a product of path integrals, each one defining the motion of a normal mode, we find that we have already solved the problem. The result is [cf. Eq. (8-10)]

$$\begin{aligned} K = \prod_{\mathbf{k}}\left(\frac{\rho kc}{2\pi i\hbar\sin kcT}\right)^{1/2}\exp\left(\frac{i\rho kc}{2\hbar\sin kcT}\left\{\left[U_1^2(\mathbf{k},T) + U_1^2(\mathbf{k},0)\right]\right.\right. \\ \left.\left.\times\cos kcT - 2U_1(\mathbf{k},T)U_1(\mathbf{k},0)\right\}\right) \end{aligned} \tag{8-131}$$

In the products over the components of $\mathbf{k}$, the x component, for example, takes on the values $2\pi n_x/L$, where n_x is an integer running from 0 to $N = L_x/d$. Here d is the spacing between atoms, and we have assumed that the solid we are studying has edge lengths L_x, L_y, and L_z. Of course, the approximation of continuity implies a zero spacing between the atoms, which means that the product is unbounded. However, we shall disregard such problems and concentrate only on the form of those terms showing dependence on initial and final coordinates. Thus, disregarding the radical which multiplies the exponential term of Eq. (8-131), we can write this equation approximately as

$$K \sim \exp \frac{i}{2\hbar} \iiint kc\{[U_1^2(\mathbf{k},T) + U_1^2(\mathbf{k},0)] \cos kcT - 2U_1(\mathbf{k},T)U_1(\mathbf{k},0)\}(\sin kcT)^{-1} \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{8-132}$$

The dependence of the amplitude on the boundary values $U_1(\mathbf{k},0)$ and $U_1(\mathbf{k},T)$ is contained in this last result. For any choice of these functions [and they, in turn, depend on $\mathbf{u}(\mathbf{r},0)$ and $\mathbf{u}(\mathbf{r},T)$, as shown by Eqs. (8-130)] the integration in Eq. (8-132) can be carried out, formally, and a final answer obtained. In this manner all questions about the quantum-mechanical behavior of the system can be answered, at least in principle.

8-8 QUANTUM FIELD THEORY

Suppose we are dealing with waves or modes which are described by continuous functions, like $\mathbf{u}(\mathbf{r},t)$, for which there is no atomic substructure or for which the wavelengths are long enough that we can neglect such a substructure. In this case we say that $\mathbf{u}(\mathbf{r},t)$ is a field, i.e., a function of each point in space. In the example we have just considered the field is the displacement field of sound. In this terminology the equations of motion are called the field equations. In the present chapter we have been dealing only with linear field equations. The lagrangians can be called the lagrangians for the field. The normal coordinates $\mathbf{U}(\mathbf{k},t)$ are the coordinates for the normal modes of the field. The description of these modes as quantum oscillators is called *quantizing the field*. The resultant theory is called quantum field theory, to distinguish it from the classical analysis of the equations.

As we have seen, almost all of the effort in quantum field theory is devoted to solving the classical equations of motion to find the normal modes, an activity completely within the realm of classical physics. The "quantization" consists then of no more than the additional remark that each of the normal modes is a quantum oscillator, with the energy levels $\hbar\omega(n + \frac{1}{2})$. Presented in this way, quantum field theory seems to be just a special consequence of the Schrödinger equation, and not an extra theory at all.

That is, or should be, the case for any situation in which the field variables (like sound displacement or pressure) are defined ultimately in terms of some combination of the basic mechanical variables. These basic variables describe the position of the particles, atoms, electrons, and also nuclei, which comprise the material carrying the field. For example, in the case of sound we assume that Schrödinger's equation describes the motion of the constituent parts, or atoms, in a crystal. Then we easily deduce that the long-wavelength sound waves obey the classical linear field equations, and we find that the modes are quantized.

In a few cases the classical equations of some field pertaining to a system are known, even though the quantum-mechanical derivation starting from Schrödinger's equation has not yet been made. For example, the equations describing the oscillations of a drop of nuclear matter have been guessed by classical analogy.[1] In such a situation it is an excellent guess that the modes of the field will turn out to be quantized oscillators if and when the complete quantum-mechanical derivation is worked out. Actually, not many such examples are left. Nearly all cases have by now been worked out.

Another type of field equation, fundamentally different from that described above, exists in quantum mechanics. An example of this type is Maxwell's set of electromagnetic equations, a set of linear field equations. These equations lead to a wave equation which is analogous to the one we developed for sound, although there are different polarization conditions. Just as an organ pipe has standing waves, or modes, so an electromagnetic field in a cavity can be described classically in terms of fundamental modes of oscillation. It is a natural inference that these oscillations are also quantized in the sense that each mode can have the energy levels $\hbar\omega n$ above the ground state, etc. This is the fundamental assumption of the quantum theory of electromagnetism. It is not a strict deduction from the Schrödinger equation for matter, because the electromagnetic field is not understood as a long-wavelength approximation of an atomic medium. Today, we do not think of any particular medium, but take the equations of Maxwell for granted. We simply assume they are to be quantized in the simple direct manner described above. We shall discuss this example in more detail in Chap. 9.

---

[1] M. S. Plesset, On the Classical Model of Nuclear Fission, *Am. J. Phys.*, vol. 9, no. 1, pp. 1–10, 1941.

The assumption of quantization for the electromagnetic fields turns out to be consistent with all experiments carried out so far, although there are some theoretical difficulties. These difficulties are associated with the extension of the scheme to modes of very short wavelengths. There are various effects which lead to diverging integrals if the integrations are carried toward zero wavelength. The corresponding difficulties do not really arise in a vibrating crystal because if we wish to carry the analysis into the very short wavelength region, where the wavelengths are comparable to the atomic spacing, we must drop the approximation of continuity. Then in the case of a crystal we find that there are only a finite number of modes in any finite volume, while in electrodynamics the number of modes in any volume is infinite.

When the various modes of a field are excited, we say there are "things" present which have different names for different cases. For sound or crystal vibrations we call them *phonons*, for the electromagnetic field *photons*, for meson field theory *mesons*, etc. Even electrons can be represented as being excitations of a field, but it is a field of a very different kind from what we have been discussing. It is called a Fermi field; the particles obey the exclusion principle, and the lagrangian is quantized not by representing it as a set of harmonic oscillators, but in a different way. Fields quantized as modes of harmonic oscillators are called Bose particles; they obey Bose or symmetric statistics. This just means that if one has two particles, one of wave number k_1 and one of k_2, there is only one state. There is no new state where the first has k_2 and the second has k_1. This is because for our field there is only one state with k_1 and k_2 each excited to the first level. It has energy $\hbar\omega_1 + \hbar\omega_2$, and it is meaningless to ask: After an exchange, which excitation is which? In the next chapter we discuss this in more detail for the case of photons of the electromagnetic field.

Problem 8-7 It is believed that neutral particles of spin zero (like neutral pions) can, when free, be represented by a field ϕ with a lagrangian

$$L = \int \frac{1}{2}\left[\left(\frac{\partial\phi}{\partial t}\right)^2 - c^2|\nabla\phi|^2 + \frac{\mu^2c^4}{\hbar^2}\,\phi\right] d^3\mathbf{r}\,dt \tag{8-133}$$

where μ is some constant. Show that this field has quantized states corresponding to waves $\exp(i\mathbf{k}\cdot\mathbf{r})$, where the energy of excitation is

$$\hbar\omega = \sqrt{\hbar^2k^2c^2 + \mu^2c^4} \tag{8-134}$$

If $\hbar\mathbf{k} = \mathbf{p}$ is considered as the momentum of each excitation the energy is

$$E = \sqrt{|\mathbf{p}|^2c^2 + \mu^2c^4} \tag{8-135}$$

This is the relativistic formula for the energy of a particle of momentum $\mathbf{p}$ and mass μ. (*Note:* For p^2 small it is approximately

$$E = \mu c^2 + \frac{p^2}{2\mu} + \cdots$$

the rest energy μc^2 plus the kinetic energy $p^2/2\mu$.)

We interpret the state of the field when the mode $\mathbf{k}_1$ is excited to the second quantum level, $\mathbf{k}_2$ to the first, etc., as the state of a system containing two particles of momentum $\hbar\mathbf{k}_1$, one with $\hbar\mathbf{k}_2$, etc. The ground state is considered the state in which no particles are present, and it is called the vacuum state. Excitation or deexcitation of the field oscillators corresponds to creation or annihilation of particles, and this is the way that such processes are represented in relativistic quantum field theory.

8-9 THE FORCED HARMONIC OSCILLATOR

In this chapter we have dealt with the simple oscillator or with systems that could be reduced to a set of such oscillators. But the oscillators have been free, not interacting with anything else. We must develop our analysis further if we wish to deal with such linear systems in interaction with other systems or driven by external forces. Examples of such systems include polyatomic molecules in varying external fields, colliding polyatomic molecules, crystals through which an electron is passing and exciting the oscillator modes, and other interactions of the modes with external fields. We shall not discuss the problem of interaction in general; instead, we use as a prototype the example of the interaction of atomic systems and charges with the electromagnetic field. We do this in the next chapter. Other cases may be analyzed by direct analogy.

These problems involve two aspects: (1) the resolution of the field into its component independent oscillators and (2) the interaction of each oscillator with external potentials or other systems. The resolution into oscillators has been exhaustively studied so far in this chapter.

To prepare the complete machinery for such problems, it remains only to analyze the behavior of a single oscillator disturbed by an external potential. We shall put these pieces together in the next chapter.

In this section we go back to the study of a single harmonic oscillator, but coupled linearly to some external potential or disturbance. The lagrangian for such a system is given by

$$L = \frac{M}{2}\dot{x}^2 - \frac{M\omega^2}{2}x^2 - \gamma(t)x \tag{8-136}$$

where $\gamma(t)$ is the external force. We assume for convenience that it is turned on only during a certain time interval T from $t = 0$ to $t = T$, so that the oscillator is free initially at $t = 0$ and finally at $t = T$. We have already completely solved this problem, obtaining the amplitude $K(b,a)$ that the oscillator goes from a point x_a at $t = 0$ to a point x_b at $t = T$, in Prob. 3-11. But for the present applications it is convenient to find as well the amplitude G_{mn} that the oscillator initially in state n is found at time T in state m. This representation is often more convenient than the coordinate representation.

In Sec. 8-1 we determined the wave functions ϕ_n for the free harmonic oscillator, and in Prob. 3-11 we evaluated the kernel describing the motion of a forced harmonic oscillator. This means that we can determine the amplitude G_{mn} by direct substitution into

$$G_{mn} = e^{(i/\hbar)E_mT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_m(x_b)K(x_b,T;x_a,0)\phi_n(x_a)\,dx_a\,dx_b \tag{8-137}$$

For the case $m = n = 0$ this integral is a gaussian somewhat lengthy to evaluate but presenting no special problems. The result is

$$G_{00} = \exp\left[-\frac{1}{2m\omega\hbar}\int_0^T\int_0^t \gamma(t)\gamma(s)e^{-i\omega(t-s)}\,ds\,dt\right] \tag{8-138}$$

If m and n are not equal to 0, then the integral is somewhat more complicated. However, we can use the same sort of trick that we used in Prob. 8-1. We shall ask for the amplitude that a forced harmonic oscillator goes from the state f to the state g, where these two states are defined in Prob. 8-1. This amplitude is

$$\begin{aligned} F(b,a) &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} G_{mn}f_m^*(b)f_n(a)e^{-iE_mT/\hbar} \\ &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} G_{mn}\left\{\exp\left[-\frac{M\omega}{4\hbar}(a^2+b^2)\right]\right\} \\ &\qquad \times \frac{a^nb^m}{\sqrt{m!n!}}\left(\frac{M\omega}{2\hbar}\right)^{(m+n)/2} e^{-i\omega T/2} \end{aligned} \tag{8-139}$$

where M is the mass of the particle [see Eq. (8-28)]. If we can work out $F(b,a)$, we can get G_{mn} by multiplying F by $\exp[(M\omega/4\hbar)(a^2 + b^2)]$

and developing the resulting expression in a power series in a and b. That is, we want first to solve

$$F(b,a) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(M\omega/2\hbar)(x_2-b)^2} K(x_2,T;x_1,0) e^{-(M\omega/2\hbar)(x_1-a)^2}\, dx_1\, dx_2 \tag{8-140}$$

where $K(x_2,T;x_1,0)$ is the kernel for a forced harmonic oscillator [Eq. (3-66)]. The variables appear only quadratically in the exponent of this integrand, so that the integration can be performed easily. Some of the resulting algebra is a little lengthy; however, eventually one finds

$$F(b,a) = \left\{\exp\left[-\frac{M\omega}{4\hbar}(a^2 + b^2 - 2abe^{-i\omega T}) + i\sqrt{\frac{M\omega}{2\hbar}}(a\beta + b\beta^* e^{-i\omega T}) - \frac{1}{2M\omega\hbar}\int_0^T \int_0^t \gamma(t)\gamma(s)e^{-i\omega(t-s)}\, ds\, dt\right]\right\} e^{-i\omega T/2} \tag{8-141}$$

where

$$\beta = \frac{1}{M\sqrt{2\omega}} \int \gamma(t) e^{-i\omega t}\, dt \tag{8-142}$$

$$\beta^* = \frac{1}{M\sqrt{2\omega}} \int \gamma(t) e^{+i\omega t}\, dt \tag{8-143}$$

The values of G_{00} can be obtained easily from Eq. (8-141) by setting $a = b = 0$. The result is the same as Eq. (8-138). Next we multiply by the exponential function, as described earlier, and find, by putting

$$x = \sqrt{\frac{M\omega}{2\hbar}}\, a \qquad y = \sqrt{\frac{M\omega}{2\hbar}}\, be^{-i\omega T}$$

that

$$\sum_{m=0} \sum_{n=0} G_{mn} \frac{x^n y^m}{\sqrt{m!}\sqrt{n!}} = [\exp(xy + i\beta x + i\beta^* y)]G_{00} \tag{8-144}$$

By expanding the right-hand side in powers of x and y and comparing terms, we obtain the final result:

$$G_{mn} = \frac{G_{00}}{\sqrt{m!n!}} \sum_{r=0}^{l} \frac{m!}{(m-r)!r!} \frac{n!}{(n-r)!r!} r!(i\beta)^{n-r}(i\beta^*)^{m-r} \tag{8-145}$$

where $l = m$ or n, whichever is the larger.

This completely solves the problem of a forced harmonic oscillator. We shall discuss it further and make use of it in the next chapter.

9

Quantum Electrodynamics

IN this chapter we shall discuss the interaction between charged particles and an electromagnetic field. We have seen one example of such an interaction in Sec. 7-6, where the electromagnetic field variables entered into the potential term of the lagrangian. The electromagnetic term introduced in that section is the vector potential **A**. Section 7-6 deals only with the motion in a definite given field. It does not tell us anything about how the field **A** arises or how it is affected by the moving particles. That is, the formulation of the problem does not contain any analysis of the dynamics of the field. Such an approach, using given potentials, is only an approximation. It is valid when these potentials arise from such large pieces of apparatus that the motion of the particle does not affect the potential.

In this chapter we shall be concerned not only with the way in which potentials affect the motion of the particle, but also with the way in which the particle affects the potentials. We shall start with the classical approach and use Maxwell's equations to describe the dynamics of the electromagnetic field. These equations express the field in terms of the charge density and current of the matter present.

We have found in preceding chapters that the quantum-mechanical laws which correspond to some classical system can be easily determined if only we can express the classical laws in the form of a least-action principle. Thus we have found that if the extremum of some action S, varied with respect to some variable q, corresponds to the classical equation of motion, then the quantum-mechanical laws are expressed as follows: The quantum-mechanical amplitude for any given situation, corresponding to the action S, is the path integral of $e^{iS/\hbar}$ integrated over all possible paths of the variable q which fit the conditions of the situation.

It is vital to our present approach that classical electrodynamics, as expressed by Maxwell's equations, can be written as a principle of least action. An action S exists which can be expressed in terms of the vector and scalar potentials $\mathbf{A}$ and ϕ. The determination of an extremum for this action, by variation of the field variables $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$, leads to a formulation of electrodynamics equivalent to Maxwell's equations. Hence, quantum electrodynamics results from the rule that the amplitude for an event is

$$K(2;1) = \int_1^2 e^{(i/\hbar)S[\mathbf{A},\phi]}\,\mathcal{D}\mathbf{A}(\mathbf{r},t)\,\mathcal{D}\phi(\mathbf{r},t) \tag{9-1}$$

where the path integral is over all values of **A** and ϕ at each point of space and time, subject to the boundary conditions at the initial and final points of the event.

9-1 CLASSICAL ELECTRODYNAMICS

Maxwell's Equations. We shall begin our study of electrodynamics from the customary classical fundamentals, i.e., from Maxwell's equations. We shall assume the magnetic permeability and dielectric constant are those for free space. Then, with $\mathbf{E}$ as the electric field vector, $\mathbf{B}$ as the magnetic field vector, c as the speed of light, $\mathbf{j}$ as the current density, and ρ as the charge density, Maxwell's equations are

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{9-2}$$

$$\nabla \times \mathbf{B} = \frac{1}{c}\left(\frac{\partial \mathbf{E}}{\partial t} + 4\pi\mathbf{j}\right) \tag{9-3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{9-4}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \tag{9-5}$$

These equations make sense only if charge is conserved, that is,

$$\nabla \cdot \mathbf{j} = -\frac{\partial \rho}{\partial t} \tag{9-6}$$

Equation (9-4) implies that $\mathbf{B}$ is the curl of some vector $\mathbf{A}$:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{9-7}$$

This relation does not fully determine $\mathbf{A}$; we still may specify its divergence. We choose

$$\nabla \cdot \mathbf{A} = 0 \tag{9-8}$$

This choice is not recommended if it is desired to keep the full relativistic four-dimensional symmetry of the equations in evidence. [It is not that the results using Eq. (9-8) are not relativistically invariant; for the results are independent of the choice of $\nabla \cdot \mathbf{A}$. It is, rather, that the invariance does not appear obvious at first glance.] In our case we shall deal with matter in the nonrelativistic approximation anyway (for we do not have a simple path integral for the Dirac equation). We wish to illustrate the properties of the quantized electromagnetic field, and the results are least cumbersome with the choice of Eq. (9-8).

Substitution into Eq. (9-5) shows that $\mathbf{E} + (1/c)\,\partial\mathbf{A}/\partial t$ has zero curl, so it must be the gradient of some potential

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \tag{9-9}$$

If there is no charge and no current density, the equations are easily solved. From Eqs. (9-2), (9-8), and (9-9) we see

$$\nabla \cdot \mathbf{E} = -\nabla^2\phi = 4\pi\rho \tag{9-10}$$

If $\rho = 0$, this means $\phi = 0$, and $\mathbf{E} = -(1/c)\,\partial\mathbf{A}/\partial t$. In Eq. (9-3) with $\mathbf{j} = 0$ this gives [*note:* $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$]

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0 \tag{9-11}$$

Thus, each component of $\mathbf{A}$ satisfies a wave equation.

If we expand $\mathbf{A}$ in running plane waves, that is,

$$\mathbf{A}(\mathbf{R},t) = \mathbf{a}_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{R}} \tag{9-12}$$

the equation for the amplitude $\mathbf{a}_k$ is $\ddot{\mathbf{a}}_{\mathbf{k}} = -c^2k^2\mathbf{a}_{\mathbf{k}}$, which implies that $\mathbf{a}_{\mathbf{k}}$ is a simple harmonic oscillator with frequency $\omega = kc$ for each component direction of $\mathbf{a}$. Actually there are only two independent transverse waves; the component of $\mathbf{a}_{\mathbf{k}}$ in the direction of $\mathbf{k}$ must be zero. That is the implication of Eq. (9-8), which can be rewritten as

$$\mathbf{k} \cdot \mathbf{a}_{\mathbf{k}} = 0 \tag{9-13}$$

Thus the field in free space is equivalent to a set of free harmonic oscillators with two transverse waves for each value of $\mathbf{k}$.

Problem 9-1 Show that $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{k}$ are mutually perpendicular for this plane-wave solution.

Solution with Charges and Currents Present. We shall again expand the solutions for $\mathbf{A}$, ϕ, and the current and charge density in plane waves, writing

$$\begin{aligned}
\mathbf{A}(\mathbf{R},t) &= \sqrt{4\pi}\,c \int \mathbf{a}_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{R}}\,\frac{d^3\mathbf{k}}{(2\pi)^3}\\
\phi(\mathbf{R},t) &= \int \phi_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{R}}\,\frac{d^3\mathbf{k}}{(2\pi)^3}\\
\mathbf{j}(\mathbf{R},t) &= \int \mathbf{j}_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{R}}\,\frac{d^3\mathbf{k}}{(2\pi)^3}\\
\rho(\mathbf{R},t) &= \int \rho_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{R}}\,\frac{d^3\mathbf{k}}{(2\pi)^3}
\end{aligned} \tag{9-14}$$

Problem 9-2 Explain why the charge density corresponding to a single charge e located at the point $\mathbf{q}(t)$ at time t is

$$\rho(x,y,z,t) = e\delta[x - q_x(t)]\delta[y - q_y(t)]\delta[z - q_z(t)] = e\delta^3[\mathbf{R} - \mathbf{q}(t)]$$

Show that

$$\rho_k = e e^{-i\mathbf{k}\cdot\mathbf{q}(t)} \tag{9-15}$$

The current density is $\mathbf{j}(\mathbf{R},t) = e\dot{\mathbf{q}}(t)\delta^3[\mathbf{R} - \mathbf{q}(t)]$. If we have a number of charges e_i located at $\mathbf{q}_i(t)$, the values of $\rho_{\mathbf{k}}$ and $\mathbf{j}_{\mathbf{k}}$ are

$$\rho_{\mathbf{k}} = \sum_i e_i e^{-i\mathbf{k}\cdot\mathbf{q}_i(t)} \qquad \mathbf{j}_{\mathbf{k}} = \sum_i e_i \dot{\mathbf{q}}_i(t) e^{-i\mathbf{k}\cdot\mathbf{q}_i(t)} \tag{9-16}$$

Here Eq. (9-13) still holds, and we may use it to simplify some expressions. The coefficient of expansion of $\mathbf{B}$ is $\mathbf{B}_{\mathbf{k}} = \sqrt{4\pi}\, ci(\mathbf{k} \times \mathbf{a}_{\mathbf{k}})$; of $\mathbf{E}$, it is $\mathbf{E}_{\mathbf{k}} = -i\mathbf{k}\phi_{\mathbf{k}} - \sqrt{4\pi}\,\dot{\mathbf{a}}_{\mathbf{k}}$. The coefficient of expansion of $\nabla \cdot \mathbf{E}$ is $i\mathbf{k} \cdot \mathbf{E}_{\mathbf{k}} = k^2\phi_k$, so we have

$$k^2\phi_k = 4\pi\rho_k \tag{9-17}$$

or $\phi_k = 4\pi\rho_k/k^2$. The function ϕ_k is completely determined in terms of the charge density ρ_k; there are no dynamic differential equations to solve, involving, for example, $\ddot{\phi}_{\mathbf{k}}$.

Problem 9-3 Prove that the relation $\phi_k = 4\pi\rho_k/k^2$ simply means that ϕ_k at any instant is the Coulomb potential from the charges at that instant, so that, for example, if ρ comes from a number of charges e_i at distances r_i from a point, the potential at the point is $\phi = \sum_i e_i/r_i$.

This is just the content of Eq. (9-10).

Equation (9-3) still remains to be solved, that is,

$$i\mathbf{k} \times \mathbf{B}_{\mathbf{k}} = \frac{1}{c}\dot{\mathbf{E}}_{\mathbf{k}} + \frac{1}{c}4\pi\mathbf{j}_{\mathbf{k}} \tag{9-18}$$

But $i\mathbf{k} \times \mathbf{B}_{\mathbf{k}} = -\sqrt{4\pi}\, c\mathbf{k} \times (\mathbf{k} \times \mathbf{a}_{\mathbf{k}}) = \sqrt{4\pi}\, ck^2\mathbf{a}_{\mathbf{k}}$ and $\dot{\mathbf{E}}_{\mathbf{k}} = -i\mathbf{k}\dot{\phi}_{\mathbf{k}} - \sqrt{4\pi}\,\ddot{\mathbf{a}}_{\mathbf{k}}$, and using Eq. (9-17) to express $\phi_{\mathbf{k}}$ as $4\pi\rho_{\mathbf{k}}/k^2$, we get

$$\ddot{\mathbf{a}}_{\mathbf{k}} + k^2c^2\mathbf{a}_{\mathbf{k}} = \sqrt{4\pi}\left(\mathbf{j}_{\mathbf{k}} - \frac{i\mathbf{k}\dot{\rho}_{\mathbf{k}}}{k^2}\right) = \sqrt{4\pi}\,\mathbf{j}'_{\mathbf{k}} \tag{9-19}$$

where we can call $\mathbf{j}'_{\mathbf{k}} = \mathbf{j}_{\mathbf{k}} - i\mathbf{k}\dot{\rho}_{\mathbf{k}}/k^2$ the transverse part of $\mathbf{j}_{\mathbf{k}}$. The law of conservation of current, expressed by Eq. (9-6), says that $\dot{\rho}_{\mathbf{k}} = -i\mathbf{k} \cdot \mathbf{j}_{\mathbf{k}}$, so

$$\mathbf{j}'_{\mathbf{k}} = \mathbf{j}_{\mathbf{k}} - \frac{\mathbf{k}(\mathbf{k} \cdot \mathbf{j}_{\mathbf{k}})}{k^2} \tag{9-20}$$

which means that $\mathbf{j}'_{\mathbf{k}}$ is $\mathbf{j}_{\mathbf{k}}$ less its component in the direction $\mathbf{k}$. Clearly, $\mathbf{k} \cdot \mathbf{j}'_{\mathbf{k}} = 0$.

We have certainly reduced Maxwell's equations to a very simple form—aside from the instantaneous Coulomb interaction between particles, we have no more than the equations for two transverse waves for each value of $\mathbf{k}$, the amplitude of each being a harmonic oscillator driven by the component of current in the corresponding direction. That is, if we choose two directions perpendicular to $\mathbf{k}$, say, 1 and 2, and call the components of $\mathbf{a}_{\mathbf{k}}$ in these directions $a_{1\mathbf{k}}$ and $a_{2\mathbf{k}}$, Maxwell's equations are

$$\ddot{a}_{1\mathbf{k}} + k^2c^2a_{1\mathbf{k}} = \sqrt{4\pi}\, j_{2\mathbf{k}} \tag{9-21}$$

$$\ddot{a}_{2\mathbf{k}} + k^2c^2a_{2\mathbf{k}} = \sqrt{4\pi}\, j_{2\mathbf{k}} \tag{9-22}$$

where $j_{1\mathbf{k}}$ and $j_{2\mathbf{k}}$ are the components of $\mathbf{j}_{\mathbf{k}}$ in these directions. (Why do we not need to say "of $\mathbf{j}'_{\mathbf{k}}$"?)

The Least-action Principle. The hypothesis of quantum electrodynamics[1] is that the oscillators defined in Eqs. (9-21) and (9-22) are quantum oscillators. To carry out the quantization, we must find the principle of least action which gives these defining equations of motion as well as the equations of motion of the particles in the field. The action is

$$S = S_1 + S_2 + S_3 \tag{9-23}$$

where

$$S_1 = \sum_i \frac{m_i}{2} \int |\dot{\mathbf{q}}_i|^2\, dt \tag{9-24}$$

is the action of all the particles, disregarding the field (if there are nonelectric forces between the particles, they are to be included in S_1),

$$\begin{aligned} S_2 &= \int \left[\rho(\mathbf{R},t)\phi(\mathbf{R},t) - \frac{1}{c}\,\mathbf{j}(\mathbf{R},t) \cdot \mathbf{A}(\mathbf{R},t) \right] d^3\mathbf{R}\, dt \\ &= \sum_i e_i \int \left\{ \phi[\mathbf{q}_i(t),t] - \frac{1}{c}\,\dot{\mathbf{q}}_i(t) \cdot \mathbf{A}[\mathbf{q}_i(t),t] \right\} dt \end{aligned} \tag{9-25}$$

[1] It should be pointed out here that some physicists use the term "quantum electrodynamics" to include electron-positron pair theory. In the present chapter we do not cover such problems. So for us, quantum electrodynamics means the quantum theory of the electromagnetic field.

is the action of interaction of field and particles, and

$$S_3 = \frac{1}{8\pi}\int (E^2 - B^2)\, d^3\mathbf{R}\, dt = \frac{1}{8\pi}\int \left[\left| -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \right|^2 - |\nabla \times \mathbf{A}|^2 \right] d^3\mathbf{R}\, dt \tag{9-26}$$

is the action of the field. The variables are $\mathbf{A}(\mathbf{R},t)$, $\phi(\mathbf{R},t)$, and $\mathbf{q}_i(t)$.

Problem 9-4 In Sec. 2-1 we discussed the mechanisms for obtaining the mechanical equations of motion from the form of the action S by obtaining the extremum S_{cl} under the condition $\delta S = 0$ for variations of coordinates, $\delta\mathbf{q}$. Show how Maxwell's equations can be derived from the action S defined in Eq. (9-23) by requiring $\delta S = 0$ for first-order variations of $\mathbf{A}$ and ϕ.

Since the dynamic equations are simplest in terms of the variables $\mathbf{a}_\mathbf{k}$, it is worthwhile to express the action in these variables. Substitution of the expansion given in Eqs. (9-14) into S_3 gives

$$\begin{aligned} S_3 &= \frac{1}{2}\int \left(\left| \dot{\mathbf{a}}_\mathbf{k} + i\mathbf{k}\frac{\phi_\mathbf{k}}{\sqrt{4\pi}} \right|^2 - c^2|\mathbf{k} \times \mathbf{a}_\mathbf{k}|^2 \right) \frac{d^3\mathbf{k}\, dt}{(2\pi)^3} \\ &= \frac{1}{2}\int \left(\phi_k{}^2 \frac{k^2}{4\pi} + \dot{\mathbf{a}}_\mathbf{k}^* \cdot \dot{\mathbf{a}}_\mathbf{k} - k^2c^2\mathbf{a}_\mathbf{k}^* \cdot \mathbf{a}_\mathbf{k} \right) \frac{d^3\mathbf{k}\, dt}{(2\pi)^3} \end{aligned} \tag{9-27}$$

and S_2 becomes

$$S_2 = \int \left(\rho_{-\mathbf{k}}\phi_\mathbf{k} - \sqrt{4\pi}\, \mathbf{j}_{-\mathbf{k}} \cdot \mathbf{a}_\mathbf{k} \right) \frac{d^3\mathbf{k}\, dt}{(2\pi)^3} \tag{9-28}$$

Upon substitution of $\phi_\mathbf{k} = 4\pi\rho_k/k^2$, the terms in $\phi_\mathbf{k}$ in S_2 and S_3 add to give

$$S_c = -\frac{4\pi}{2}\int \frac{\rho_\mathbf{k}\rho_{-\mathbf{k}}}{k^2}\frac{d^3\mathbf{k}}{(2\pi)^3} = -\frac{1}{2}\sum_i\sum_j \frac{e_ie_j}{|\mathbf{q}_i - \mathbf{q}_j|} \tag{9-29}$$

using Eqs. (9-16), since $\int (4\pi/k^2)[\exp(i\mathbf{k}\cdot\mathbf{R})]\, d^3\mathbf{k} = 1/R$. This is just the Coulomb interaction between the charges, which is usually considered in analyzing atoms when electromagnetic radiation effects are neglected. That is, we shall include this interaction in the action S_{mat} of the matter,

$$S_{mat} = S_1 + S_c = \int \sum_i \left(\frac{m_i}{2}\dot{q}_i{}^2 - \frac{1}{2}\sum_j \frac{e_ie_j}{|\mathbf{q}_i - \mathbf{q}_j|} \right) \tag{9-30}$$

and write $S = S_{mat} + S_{int} + S_{rad}$. We have thus divided the actions of the electromagnetic field S_3 into two parts; one contributes to an instantaneous Coulomb interaction, and the remainder we shall call the radiation field S_{rad}. (The radiation field takes care of all corrections to the instantaneous field, such as, for example, that the total effects are retarded and act no faster than the speed of light.) The action of the radiation field is S_3 less the terms involving $\phi_{\mathbf{k}}$. That is,

$$S_{rad} = \frac{1}{2}\int \left(\dot{a}^*_{1\mathbf{k}}\dot{a}_{1\mathbf{k}} - k^2c^2a^*_{1\mathbf{k}}a_{1\mathbf{k}} + \dot{a}^*_{2k}\dot{a}_{2k} - k^2c^2a^*_{2k}a_{2k}\right)\frac{d^3\mathbf{k}\,dt}{(2\pi)^3} \tag{9-31}$$

which is just the action of the radiation oscillators. The action of interactions of these oscillators with the particles is

$$S_{int} = \sqrt{4\pi}\int \left(j_{1-\mathbf{k}}a_{1\mathbf{k}} + j_{2-\mathbf{k}}a_{2\mathbf{k}}\right)\frac{d^3\mathbf{k}\,dt}{(2\pi)^3} \tag{9-32}$$

Clearly, the variation of the total action of S with respect to the $a_{1\mathbf{k}}$ and $a_{2\mathbf{k}}$ gives the equations of motion (9-21) and (9-22).

Written more explicitly, the action S_{int} is

$$S_{int} = \sqrt{4\pi}\sum_j \int \left(a_{1\mathbf{k}}\dot{q}_{1j} + a_{2\mathbf{k}}\dot{q}_{2j}\right)e^{i\mathbf{k}\cdot\mathbf{q}_j(t)}\frac{d^3\mathbf{k}\,dt}{(2\pi)^3} \tag{9-33}$$

where q_{1j} and q_{2j} are the components of $\mathbf{q}_j$ in the direction transverse to $\mathbf{k}$. Thus all the laws of nonrelativistic mechanics and of electrodynamics are contained in the proposition that S, the sum of Eqs. (9-30), (9-31), and (9-33), is stationary for variations in the paths of the variables $\mathbf{q}_j(t)$, $a_{1\mathbf{k}}(t)$, and $a_{2\mathbf{k}}(t)$. Quantum electrodynamics results from integrating $e^{i\hbar/S}$ over these paths, and it is described in the next section.

9-2 THE QUANTUM MECHANICS OF THE RADIATION FIELD

We begin by discussing the quantum mechanics of the radiation field in empty space. There is no matter present, so the total action is that of the radiation field along

$$S = S_{rad} \tag{9-34}$$

as given by Eq. (9-31). It is evidently the action of a set of harmonic oscillators. We have seen some examples of expressions like Eq. (9-31) in Chap. 8. We make the assumption discussed in Chap. 8 that quantum electrodynamics results from considering these as quantum-mechanical oscillators.

The modes of our system are running waves, two for each value of $\mathbf{k}$

(polarization 1 and 2) with frequency $\omega = kc$. For one of the modes, say, $a_{1\mathbf{k}}$, the available energy levels are

$$E_{1\mathbf{k}} = (n_{1\mathbf{k}} + \tfrac{1}{2})\hbar kc \tag{9-35}$$

where $n_{1\mathbf{k}}$ is any positive integer or zero.

If $n_{1\mathbf{k}} = 1$, we say there is one photon present of polarization 1 and momentum $\hbar k$; in general, we say $n_{1\mathbf{k}}$ such photons are present. The energy of a single photon of this kind is $\hbar kc$.

Problem 9-5 The momentum in the field is given by $(1/4\pi c) \times \int \mathbf{E} \times \mathbf{B}\, d(\text{vol})$. In the absence of matter (so $\phi_{\mathbf{k}} = 0$), show this is $\int \mathbf{k}(\mathbf{a}_{\mathbf{k}}^* \cdot \dot{\mathbf{a}}_{\mathbf{k}})\, d^3\mathbf{k}/(2\pi)^3$.

Later on when we consider the interaction of matter with the radiation field, we shall find that the matter absorbs or emits one photon at a time of energy $\hbar\omega$. Of course, this is the same as Planck's original hypothesis.

It is quite striking and surprising that the states n of the oscillators can also be described by imagining that there are n "particles" or "photons" present. It is clear, of course, that the energy values agree. But there is one further subtle point that must be noted before the oscillator states can be completely successfully described as particles. Suppose, for example, that just two of the n_j differ from zero, say, $n_a = 1$, $n_b = 1$. This single state we may wish to represent by saying that we have one photon in state a and another in state b. But at first sight this way of speaking might seem to imply that there were two states available, both of the same energy. For we could also expect to be able to put the first photon in state b and the second in state a. The way out of this can be seen when we consider the example of α particles. Suppose we have two α particles with coordinates x and y, and say the x particle is in a state represented by $f(x)$ and the y particle is in a state $g(y)$. Thus the wave function for the system would be

$$\psi(x,y) = f(x)g(y) \tag{9-36}$$

a function of the two variables x, y. But another state might have y in the state f and x in g, leading to another state of wave function

$$\psi(x,y) = g(x)f(y) \tag{9-37}$$

which differs from the first. But if the particles are truly identical, like α particles, the two states are indistinguishable. As we described in Chap. 1, it turns out to be a rule of quantum mechanics (not derivable from the Schrödinger equation) that for α particles the amplitudes for two cases which differ only by exchange of the α particles must always be added. The only allowed wave function is in this case

$$\psi(x,y) = f(x)g(y) + g(x)f(y) \tag{9-38}$$

(suitably normalized: if f and g are orthonormal, the factor is $1/\sqrt{2}$; if $f = g$ and they are normalized, it is $\frac{1}{2}$). In general $\psi(x,y) = \psi(y,x)$ for α particles and for other particles obeying Bose statistics. There is, for such particles, only one state: one particle in state f, the other in state g.

It turns out that all the results are consistent if, when we consider oscillator excitation states as representing numbers of photons, we also say that photons are Bose particles. Then the single state $n_a = 1$, $n_b = 1$ represents the situation that there are two photons, one in a, one in b. Exchange does not produce a new state.

For electrons of parallel spin or other Fermi particles we must subtract the amplitudes when the identity of the particles is: reversed

$$\psi(x,y) = f(x)g(y) - g(x)f(y) \tag{9-39}$$

The wave function $\psi(x,y) = -\psi(y,x)$ is antisymmetric in general for Fermi particles. This is, of course, also just one state. But for Fermi particles, two particles cannot occupy the same state. If we put $f = g$ into Eq. (9-39), we get zero. Two photons, like two α particles, can occupy the same state; for photons it corresponds to the $n = 2$ oscillator states.

There is one particular situation with matter present which, in the ideal case, can be handled nearly as simply as the matter-free case. That is the case of a cavity resonator (or a wave guide) where the walls may be idealized as perfect conductors. Then classically, as is well known, there are a number of possible oscillator modes with more or less complicated distributions of electric fields. The classical action is then reducible to a set of free oscillators, but the variables now represent the amplitudes of the various modes, rather than the amplitudes of plane running waves. These oscillators are then analyzed as quantum oscillators, and we speak of the number of photons in each mode.

9-3 THE GROUND STATE

Vacuum Energy. The state of the electromagnetic field of lowest possible energy, which we shall call the ground state or the vacuum state, is that in which there are no photons in any mode. This means that the energy in each mode is $\hbar\omega/2$, where ω is the frequency of the mode. Now if we were to sum this ground-state energy over all of the infinite number of possible modes of ever-increasing frequency which exist even for a finite box, the answer would be infinity. This is the first symptom of the difficulties which beset quantum electrodynamics.

In the present case, for the vacuum state, the trouble is easily fixed. Suppose we choose to measure energy from a different zero point. Since there is no physical effect resulting from a constant energy, the result of any experiment we perform will be insensitive to the arbitrary choice of the zero point in energy. Therefore, we assign to the vacuum state the energy zero. Then the total energy in any state of the electromagnetic field is given by

$$E = \sum_j n_j \hbar\omega_j \tag{9-40}$$

where the sum is taken over all the modes j of the field.

Unfortunately, it is really not true that the zero point of energy can be assigned completely arbitrarily. Energy is equivalent to mass, and mass has a gravitational effect. Even light has a gravitational effect, for light is deflected by the sun. So, if the law that action equals reaction has qualitative validity, then the sun must be attracted by the light. This means that a photon of energy $\hbar\omega$ has a gravity-producing effect, and the question is: Does the ground-state energy term $\hbar\omega/2$ also have an effect? The question stated physically is: Does a vacuum act like a uniform density of mass in producing a gravitational field?

Since most of the space is a vacuum, any effect of the vacuum-state energy of the electromagnetic field would be large. We can estimate its magnitude. First, it should be pointed out that some other infinities occurring in quantum-electrodynamic problems are avoided by a particular assumption called the *cutoff rule.* This rule states that those modes having very high frequencies (short wavelength) are to be excluded from consideration. The rule is justified on the ground that we have no evidence that the laws of electrodynamics are obeyed for wavelengths shorter than any which have yet been observed. In fact, there is a good reason to believe that the laws cannot be extended to the short-wavelength region. Mathematical representations which work quite well at longer wavelengths lead to divergences if extended into the short-wavelength region. The wavelengths in question are of the order of the Compton wavelength of the proton; $1/2\pi$ times this wavelength is $\hbar/Mc \simeq 2 \times 10^{-14}$ cm.

For our present estimate suppose we carry out sums over wave numbers only up to the limiting value $k_{\text{max}} = Mc/\hbar$. Approximating the sum over states by an integral, we have, for the vacuum-state energy per unit volume,

$$\frac{E_0}{\text{unit vol}} = 2\,\frac{\hbar c}{2(2\pi)^3}\int_0^{k_{\text{max}}} k \cdot 4\pi k^2\,dk = \frac{\hbar c k_{\text{max}}^4}{8\pi^2} \tag{9-41}$$

(Note the first factor 2, for there are two modes for each k.) The equivalent mass of this energy is obtained by dividing the result by c^2. This gives

$$\frac{m_0}{\text{unit vol}} = 2.10^{15}\ \text{g/cm}^3 \tag{9-42}$$

Such a mass density would, at first sight at least, be expected to produce very large gravitational effects which are not observed. It is possible that we are calculating in a naive manner, and, if all of the consequences of the general theory of relativity (such as the gravitational effects produced by the large stresses implied here) were included, the effects might cancel out; but nobody has worked all this out. It is possible that some cutoff procedure that not only yields a finite energy density for the vacuum state but also provides relativistic invariance may be found. The implications of such a result are at present completely unknown.

For the present we are safe in assigning the value zero to the vacuum-state energy density. Up to the present time no experiments that would contradict this assumption have been performed. As we progress further into the field of quantum electrodynamics we shall find other divergent integrals which are more difficult to circumvent.

Vacuum Wave Function. The wave function for the set of oscillators is just the product of the wave functions for each mode. For the ground state the wave function of the oscillator 1, **k** is proportional to exp $[-(ck/2\hbar)/a^*_{1\mathbf{k}}a_{1\mathbf{k}}]$, so the wave function for the entire system in the ground state, or vacuum state, is, within a normalization constant,

$$\Phi_0 = \exp\left[-\sum_{\mathbf{k}} \frac{kc}{2\hbar}\,(a^*_{1\mathbf{k}}a_{1\mathbf{k}} + a^*_{2\mathbf{k}}a_{2\mathbf{k}})\right] \tag{9-43}$$

Problem 9-6 Show, using sine and cosine modes and real variables, that this expression using complex variables is indeed correct (cf. Prob. 8-4).

Problem 9-7 Show, for the vacuum state, the mean value of $a^*_{1\mathbf{k}}a_{1\mathbf{l}}$ is $(\hbar/2kc)\delta_{kl}$. Develop a formula for the mean of $(a^*_{1\mathbf{k}}a_{1\mathbf{k}})^r$ for integral r and explain thereby how the mean of such quantities as $(a^*_{1\mathbf{k}}a_{1\mathbf{k}})^r(a^*_{1\mathbf{p}}a_{1\mathbf{p}})^s$ can be got for $\mathbf{p} \neq \mathbf{k}$. Show that the mean of $(a_{1\mathbf{k}})^2$ or $(a^*_{1\mathbf{k}})^2$ vanishes. Show that the mean of the product of any odd number of a's is zero and that you can compute the expectation value of any product of a's or a^*'s for the vacuum state.

Problem 9-8 For the state for which there is just one photon present in state $1\mathbf{k}$, all of the factors in the wave function are ϕ_0 except one, which is ϕ_1. But for an oscillator $\phi_1(q) = q\phi_0(q)$. The wave function representing an excited running wave is a linear superposition of the state with the cosine mode excited and i times the state with the sine wave excited, so show that the wave function for just one photon present in state $1\mathbf{k}$ is $a^*_{1\mathbf{k}}\Phi_0$. This is not normalized. The normalization is $\int \Phi^*_0 a_{1\mathbf{k}} a^*_{1\mathbf{k}} \Phi_0$, or the expectation of $a_{1\mathbf{k}}a^*_{1\mathbf{k}}$ for the vacuum, which we have seen in the preceding problem is $\hbar/2kc$. Hence the normalized one-photon state is $\sqrt{2kc/\hbar}\, a^*_{1\mathbf{k}}\Phi_0$.

9-4 INTERACTION OF FIELD AND MATTER

To deal with the interaction of the radiation field with matter is not difficult in a formal way. Evidently from the action expression of Eqs. (9-30), (9-31), and (9-33) we see we must deal with the matter system interacting with the radiation oscillators and must calculate amplitudes from

$$\text{Amplitude} = \int \left\{ \exp\left[\frac{i}{\hbar}\left(S_{mat} + S_{int} + S_{rad}\right)\right]\right\} \prod_{i\mathbf{k}} \mathfrak{D}\mathbf{q}_i\, \mathfrak{D}a_{1\mathbf{k}}\, \mathfrak{D}a_{2\mathbf{k}} \tag{9-44}$$

The coordinates of the radiation oscillators can be integrated out immediately; for they appear only in quadratic expressions. We shall do this integration in the next section.

Emission from an Atom. Part of the complication in this problem is simply the confusion produced by so many coordinates and states. So we shall begin by dealing with a simple problem just to get more used to what is involved. We shall solve the problem of the probability of emission of light by a single atom, using perturbation theory (assuming the interaction S_{int} of light and matter is small and expanding it only to the first order).

If S_{int} is neglected, the radiation and matter are independent systems. Let the states of the atom alone have energy e_N for various values of N with wave functions $\Psi_N(\mathbf{q})$, where $\mathbf{q}$ represents the $\mathbf{q}_i$ of all the particles of the atom. The state of the radiation can be defined by giving the values of all the integers n_{1k} and n_{2k}. The energy levels of the combined system are

$$E = e_N + \sum_{\mathbf{k}} (n_{1\mathbf{k}} + n_{2\mathbf{k}})\hbar kc \tag{9-45}$$

The wave function for this state is a product

$$\Psi = \psi_N(\mathrm{q})\Phi(n_{1\mathbf{k}},n_{2\mathbf{k}}) \tag{9-46}$$

where $\Phi(n_{1\mathbf{k}},n_{2\mathbf{k}})$ is the wave function for the radiation field (a product of harmonic oscillator wave functions).

To deal with atomic radiation of a photon, we consider as the initial state that the atom is in some level M and no photons are present (all $n_{1\mathbf{k}}$ and $n_{2\mathbf{k}}$ equal 0). This wave function is

$$\Psi_i = \psi_M(\mathrm{q})\Phi_0 \tag{9-47}$$

with Φ_0 from Eq. (9-43). In the final state the atom is in another level N, but now a photon is present, say of momentum $\boldsymbol{l}$ and polarization 1. According to Prob. 9-8 the wave function of the radiation alone is $a^*_{1\boldsymbol{l}}\Phi_0$; the complete final wave function is

$$\Psi_f = \sqrt{\frac{2lc}{\hbar}}\,\Psi_N(\mathrm{q})a^*_{1\boldsymbol{l}}\Phi_0 \tag{9-48}$$

Now to find the transition probability per second (to first order) we see, according to Eq. (6-79), we shall need the matrix element V_{fi} of the perturbation potential between these states. The perturbation action is S_{int}, as defined in Eq. (9-32), and the corresponding potential is

$$V = \sqrt{4\pi}\sum_{\mathbf{k}} (a^*_{1\mathbf{k}}j_{1\mathbf{k}} + a_{1\mathbf{k}}j^*_{1\mathbf{k}}) \tag{9-49}$$

where $j_{1\mathbf{k}}$ depends on the atomic variables, as in Prob. 9-2. This matrix element is

$$V_{fi} = \int \psi^*_N\Phi^*_0 \sqrt{\frac{2lc}{\hbar}}\,a_{1\boldsymbol{l}} \sum_{\mathbf{k}} \sqrt{4\pi}\,(a^*_{1\mathbf{k}}j_{1\mathbf{k}} + a_{1\mathbf{k}}j^*_{1\mathbf{k}})\psi_M\Phi_0\,dq \prod_{\mathbf{k}} da_{1\mathbf{k}} \tag{9-50}$$

which gives

$$\begin{aligned} V_{fi} = {} & \sum_{\mathbf{k}} \sqrt{\frac{8\pi lc}{\hbar}} \int \Phi^*_0 a_{1\boldsymbol{l}}a^*_{1\mathbf{k}}\Phi_0 \prod_{k} da_{1\mathbf{k}} \int \psi^*_N j_{1\mathbf{k}}\psi_M\,dq \\ & + \sum_{\mathbf{k}} \int \sqrt{\frac{8\pi lc}{\hbar}}\,\Phi^*_0 a_{1\boldsymbol{l}}a_{1\mathbf{k}}\Phi_0 \prod_{\mathbf{k}} da_{1\mathbf{k}} \int \psi^*_N j^*_{1\mathbf{k}}\psi_M\,dq \end{aligned} \tag{9-51}$$

because only j depends on q. The expectation values of the a's for the vacuum state were worked out in Prob. 9-7, that is,

$$\int \Phi_0^* a_{1l} a_{1\mathbf{k}}^* \Phi_0 \prod_{\mathbf{k}} da_{1\mathbf{k}} = 0$$

unless $k = l$, in which case it is $\hbar/2lc$. Let us write the matrix element $\int\psi_N^*\mathbf{j}\psi_M$ as $(\mathbf{j})_{NM}$. Our matrix element is therefore $\sqrt{2\pi\hbar/lc}\,(j_{1l})_{NM}$. The probability of transition per second is then [cf. Eq. (6-94)]

$$\left(\frac{2\pi}{\hbar}\right)\left(\frac{2\pi\hbar}{lc}\right)|j_{1l}|_{NM}{}^2\delta(E_M - E_N - hlc) \tag{9-52}$$

Ordinarily we are not interested in the problem of exciting one particular photon but would rather see the probability of emission of any photon (of the polarization 1) into some small solid angle $d\Omega$. We must sum $\mathbf{l}$ over all values which correspond to this direction. The number of values of $\mathbf{l}$ per unit volume is $d^3\mathbf{l}/(2\pi)^3$, or if $\mathbf{l}$ is in the specified direction, we require the integral over dl of $l^2\,dl\,d\Omega/(2\pi)^3$, so that we find that the probability of a transition per second is

$$\frac{dP}{dt} = \int \frac{(2\pi)^2}{lc}|j_{1l}|_{NM}{}^2\delta(E_M - E_N - \hbar lc)l^2\frac{dl\,d\Omega}{(2\pi)^3} \tag{9-53}$$

The integral on l gives

$$\frac{dP}{dt} = \frac{\omega}{2\pi\hbar c^3}|j_{1l}|_{NM}{}^2\,d\Omega \tag{9-54}$$

for the rate of emission of light of polarization 1 in direction $\mathbf{l}$ into the solid angle $d\Omega$. The frequency emitted satisfied

$$\omega = lc = \frac{E_M - E_N}{\hbar} \tag{9-55}$$

Problem 9-9 For a complicated system moving nonrelativistically

$$(j_{1\mathbf{k}})_{NM} = \sum_b (e_b\mathbf{e}\cdot\dot{\mathbf{q}}_b e^{-i\mathbf{k}\cdot\mathbf{q}b})_{NM} \tag{9-56}$$

where $\mathbf{e}$ is a unit vector in the direction of the polarization of the light and e_b and $\mathbf{q}_b$ are the charge and position of the bth particle. Assume the wavelength of the light is very large compared with the size of the atom, i.e., that the absolute square of the wave function describing the position of the bth electron falls to 0 over a distance small compared with $1/k$. Show that we can then approximate $\exp\,(i\mathbf{k}\cdot\mathbf{q}_b/\hbar)$ by unity and write the matrix element as

$$j_{1\mathbf{k},NM} = i\omega \mathbf{e} \cdot \boldsymbol{\mu}_{NM} \tag{9-57}$$

where

$$\boldsymbol{\mu}_{NM} = \sum_b (e_b \mathbf{q}_b)_{NM} \tag{9-58}$$

The function $\boldsymbol{\mu}_{NM}$ is called the *matrix element of the electric dipole moment* of the atom, and the approximation used to derive Eq. (9-57) is called the *dipole approximation.* Show that the total probability to emit light in any direction per unit time is

$$\frac{dP}{dt} = \frac{4\omega^3}{3\hbar c^3} |\boldsymbol{\mu}_{NM}|^2 \tag{9-59}$$

[Integrate Eq. (9-54) over all directions, remembering that **e** is perpendicular to **k** and that there are two possible directions of polarization.]

Elimination of Electromagnetic Field Variables. Since the radiation field is represented by a quadratic action functional, we can integrate out all its coordinates. We shall do so here. We must integrate all the variables $a_{1\mathbf{k}}$, $a_{2\mathbf{k}}$ in Eq. (9-44). We must specify the initial and final states of the radiation field. First we shall take the simplest case that initially and finally we have a vacuum, the oscillators all going from 0 to 0 photon number. Our amplitude can be written

$$\text{Amplitude} = \int e^{(i/\hbar)S_{mat}} X[q]\, \mathfrak{D}q \tag{9-60}$$

where

$$X[q] = \int e^{(i/\hbar)(S_{int}+S_{rad})} \prod_{\mathbf{k}} da_{1\mathbf{k}}\, da_{2\mathbf{k}} \tag{9-61}$$

is a functional of the q's which appear on the right-hand side of the equation in the current variables j. Since the action is a sum of contributions $\sum_k (S_{1\mathbf{k}} + S_{2\mathbf{k}})$ from each mode, where

$$S = \int \left[\sqrt{4\pi}\, (ja^* + j^*a) + \tfrac{1}{2}\dot{a}^*\dot{a} - \frac{k^2c^2}{2}\, a^*a - \frac{\hbar kc}{2} \right] dt \tag{9-62}$$

clearly X is a product of corresponding factors. The integral for one typical mode,

$$X_{1\mathbf{k}} = \int \left\{ \exp \left[\frac{i}{\hbar} \int \left(\sqrt{4\pi}\, j^*_{1\mathbf{k}} a_{1\mathbf{k}} + \sqrt{4\pi}\, j_{1\mathbf{k}} a^*_{1\mathbf{k}} + \tfrac{1}{2} \dot{a}^*_{1\mathbf{k}} \dot{a}_{1\mathbf{k}} - \frac{k^2c^2}{2} a^*_{1\mathbf{k}} a_{1\mathbf{k}} - \frac{\hbar kc}{2} \right) dt \right] \right\} \mathfrak{D}a_{1\mathbf{k}}$$

$$= \exp \left[-\frac{4\pi}{2\hbar} \int j_{1\mathbf{k}}(t) j^*_{1\mathbf{k}}(s) \frac{1}{2kc} e^{-ikc|t-s|}\, dt\, ds \right] \tag{9-63}$$

is a type of path integral which we have already done many times, except for the complication of complex variables, which can first be reduced to real variables. In fact, this is exactly the problem discussed in Sec. 8-9. The interaction function $\gamma(t)$ of Eq. (8-136) is here given by $\gamma = \sqrt{4\pi}\, j_{1\mathbf{k}}$ and $\omega = kc$. The final expression of Eq. (9-63) is equivalent to Eq. (8-138). The product of such factors for each k and polarization gives $X = e^{iI/\hbar}$, where

$$I = \frac{1}{2} \sum_{\mathbf{k}} \int [j_{1\mathbf{k}}(t) j^*_{1\mathbf{k}}(s) + j_{2\mathbf{k}}(t) j^*_{2\mathbf{k}}(s)] \frac{4\pi}{2kc} e^{-ikc|t-s|}\, dt\, ds \tag{9-64}$$

Thus the problem of a vacuum-to-vacuum transition is completely solved in terms of a path integral on the matter variables alone:

$$\text{Amplitude} = \int e^{(i/\hbar)(S_{mat}+I)}\, \mathfrak{D}q(t) \tag{9-65}$$

We shall discuss a number of consequences of this result. (The case that the initial or final state is not a vacuum is described in the next section.)

It appears that the net result is simply this: The matter acts not with S_{mat} but with a modified action $S'_{mat} = S_{mat} + I$. The modification results from a reaction with the electromagnetic field. This is not true in a strictly classical sense, for the action I is a complex number. It can be shown that the classical physics which results from using the principle of least action, with the real part of S'_{mat} only, is exactly equivalent to the combination of Maxwell's equations and Newton's laws. But it does not correspond to the case that Maxwell's equations are solved by using just retarded waves. (In fact, a restriction to retarded waves cannot be represented by any principle of least action in which only matter coordinates appear. Instead it corresponds to using half the advanced and half the retarded solution.[1]) Our full quantum-mechanical complex expression for I is correct, and we shall now look at its consequences.

[1] J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.*, vol. 17, p. 157, 1945.

First-order Perturbation Expansion. The integral on q is too complicated to do exactly, but in the expression for the currents in I the charge e of the particles occurs. Thus I is proportional to e^2, which in dimensionless form is the fine-structure constant

$$\frac{e^2}{\hbar c} = \frac{1}{137.039}$$

a small number whose exact value and meaning are unknown except experimentally. Thus we may expect that the effect of I is small. We already know that the Schrödinger theory gives atomic levels, for example, quite accurately. There can be only small errors arising from the neglect of I. Let us look at the effect of I in first order in e^2, corresponding to second order in e on the original action of Eq. (9-32). Let us take the transition amplitude λ_{MM} as defined in Sec. 6-5, where the matter system begins and ends in state M. If I is neglected, the zero order is

$$\lambda_{MM}{}^0 = e^{-(i/\hbar)E_M t} \tag{9-66}$$

The first-order term is

$$\begin{aligned}\lambda_{MM}{}^1 &= \frac{1}{\hbar}\int_{t_i}^{t_f} \psi_M^*(q_f)e^{(i/\hbar)S_{mat}} I \psi_M(q_i)\,\mathcal{D}q(t)\\ &= -\frac{1}{2\hbar}\sum_{\mathbf{k}}\int \psi_M^*(q_f)e^{(i/\hbar)S_{mat}}\int [j_{1\mathbf{k}}^*(t)j_{1\mathbf{k}}(s)\\ &\qquad + j_{2\mathbf{k}}^*(t)j_{2\mathbf{k}}(s)]\,\frac{4\pi}{2kc}\,e^{-ikc|t-s|}\,dt\,ds\,\psi_M(q_i)\,\mathcal{D}q(t)\end{aligned} \tag{9-67}$$

Now we take $t > s$ and double the result. The evaluation of a similar expression was worked out in Sec. 5-1. For the present example, we get

$$\lambda_{MM}{}^1 = -\frac{i}{\hbar}(\Delta E)Te^{-iE_M T/\hbar}$$

where

$$\begin{aligned}\Delta E &= i\sum_N\sum_{\mathbf{k}}\frac{4\pi}{2kc}[(j_{1\mathbf{k}}^*)_{MN}(j_{1\mathbf{k}})_{NM} + (j_{2\mathbf{k}}^*)_{MN}(j_{2\mathbf{k}})_{NM}]\\ &\quad\times\int_0^\infty e^{(i/\hbar)(E_M-E_N-\hbar kc)\tau}\,d\tau = \sum_N 4\pi\hbar\int [|(j_{\mathbf{k}})_{NM}|^2\\ &\qquad + |(j_{2\mathbf{k}})_{NM}|^2][2kc(E_M - E_N - k\hbar c + i\epsilon)]^{-1}\frac{d^3\mathbf{k}}{(2\pi)^3}\end{aligned} \tag{9-68}$$

This has a real and imaginary part and can be written as

$$\Delta E = \delta E - \frac{i\hbar\gamma}{2}$$

The real part δE represents a small shift in the energy levels of the atom, called the Lamb shift. Such a shift was discovered experimentally by Lamb and Rutherford. This is

$$\delta E = \sum_N \int [|(j_{1\mathbf{k}})_{NM}|^2 + |(j_{2\mathbf{k}})_{NM}|^2] \text{P.P.}(E_M - E_N - \hbar kc)^{-1} \frac{4\pi\hbar}{2kc} \frac{d^3\mathbf{k}}{(2\pi)^3} \qquad (9\text{-}69)$$

and the imaginary part is

$$\frac{\hbar\gamma}{2} = \sum_N \int [|(j_{1\mathbf{k}})_{NM}|^2 + |(j_{2\mathbf{k}})_{NM}|^2] \pi\delta(E_M - E_N - \hbar kc) \frac{4\pi\hbar}{2kc} \frac{d^3\mathbf{k}}{(2\pi)^3} \qquad (9\text{-}70)$$

The amplitude that the atom remains in the upper state with no photons emitted goes as $\exp[-i(E_m + \delta E - i\gamma/2)T/\hbar]$ and the probability as $e^{-\gamma T}$. That is, the probability to remain in the state M decreases exponentially with the rate decay γ. Physically it should decrease because the atom in the state M can emit a photon and fall to a lower state N. Comparison with Eq. (9-53) shows that γ in Eq. (9-70) is indeed the total rate of transition from state M to all lower states N.

9-5 A SINGLE ELECTRON IN A RADIATIVE FIELD

The Energy Correction. In order to study the electromagnetic energy correction δE, we shall consider the simplest case: that in which the matter system has only one moving charge (e.g., a hydrogen atom with an infinitely heavy nucleus or a free electron in empty space) whose coordinates we shall call $\mathbf{R}$. Thus $\mathbf{j} = e\dot{\mathbf{R}} \exp(i\mathbf{k} \cdot \mathbf{R}/\hbar)$. We have here a case where $\mathbf{j}$ contains $\dot{\mathbf{R}}$, and in considering second-order terms we must take appropriate care, as discussed in Sec. 7-3. There is an additional term to δE from the squared velocity term $\dot{R}^2$. Expressing $\dot{\mathbf{R}}$ in terms of the momentum operator $\mathbf{p}$, as in Sec. 7-5, we obtain

$$\delta E_1 = \sum_N \int \frac{d^3\mathbf{k}}{2kc(2\pi)^3} \left(|\mathbf{p}_1 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM} + |\mathbf{p}_2 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM}\right) \frac{4\pi e^2\hbar}{m^2(E_M - E_N - \hbar kc)} + \frac{4\pi e^2}{m} \int \frac{\hbar}{2kc} \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{9-71}$$

Problem 9-10 Why do we not need to be careful to write $\frac{1}{2}[\mathbf{p}_1 \exp(-i\mathbf{k}\cdot\mathbf{R}/\hbar) + \exp(-i\mathbf{k}\cdot\mathbf{R}/\hbar)\mathbf{p}_1]$ in the matrix elements?

Let us take the simplest case of a free electron at rest. Any δE_R we get for energy in the field will represent a correction to the rest energy, or as can be shown from the relativity theory, to the mass, $\delta m = \delta E_R/c^2$. This is the so-called *electromagnetic mass correction.* For a free particle at rest, the states are plane waves. If the momenta in M and N are $\mathbf{p}_M$ and $\mathbf{p}_N$, the matrix element $[\mathbf{p}_1 \exp(-i\mathbf{k}\cdot\mathbf{R}/\hbar)]_{NM}$ is zero unless $\mathbf{p}_N = \mathbf{p}_M - \mathbf{k}$, in which case it is $\mathbf{p}_{1N}$. Thus for an electron at rest initially, the matrix element is 0 and δE_R is just the last integral of Eq. (9-71), which is infinite!

Short-wavelength Difficulties. This is not the whole of it. When we eliminated the term $\rho_k\rho_k/2k^2$ in S_c, we pointed out that this represented the interaction between point charges $\frac{1}{2}\sum_{i,j} e_i e_j |\mathbf{q}_i - \mathbf{q}_j|^{-1}$, but neglected to point out that the infinite terms $i = j$ must also be included in the sum. Indeed, for a single particle $\rho_k = e\exp(i\mathbf{k}\cdot\mathbf{q})$, so $\frac{1}{2}|\rho_k|^2/k^2 = 4\pi e^2/2k^2$ and the term is $4\pi e^2 \int (\frac{1}{2}k^2)\, d^3k/(2\pi)^3$. The infinities here and in δE_R above do not cancel, and we are left with a real difficulty; our integrals over momentum k diverge quadratically. Quantum electrodynamics gives nonsensical results.

It is true that we are using a nonrelativistic treatment of the charged particle. The relativistic treatment of the matter (quantum electrodynamics is not altered) does not rid us of the divergent results, although the order of infinity may be changed. For a particle of spin 0, like a π meson, the order is unchanged; it is still a quadratic divergence. Here there is presumably an experimental value of the mass correction available. As far as is known, through other interactions, the sole difference of charged and neutral π mesons is the charge, i.e., the different way they couple to the electromagnetic field. So presumably the mass difference of the charged π meson with a mass m_π of 273.2 electron masses and the neutral π mesons of 264.2

electron masses, that is, 9.0 electron masses = $0.034m_\pi$, or 4.6 Mev, represents energy in the electromagnetic field.

If we arbitrarily stop our integrals at some higher momentum $k_{\max}$ (which is not a relativistically invariant procedure), we get an energy $e^2(k_{\max})^2/2\pi m_\pi c$ from the last term of Eq. (9-71), which is the largest term if $\hbar k_{\max}/c$ is very much larger than the pion mass m_π. If this equals $0.034m_\pi c^2$, then $(e^2/2\pi\hbar c)(\hbar k_{\max}/m_\pi c)^2 = 0.034$, or

$$k_{\max} \approx \frac{5.4m_\pi c}{\hbar} \approx \frac{0.8Mc}{\hbar}$$

where M is the mass of a proton. (The relativistic theory gives $\Delta E = 0.034m_\pi$ with a cutoff at about the same energy.) It is for this reason that we conclude that our present-day formulation of quantum electrodynamics (or of the "particles" with which photons interact) is faulty. The fault lies in the way we deal with energies beyond proton mass or with corresponding frequencies, or wave numbers. The difficulty arises with modes whose wavelength is less than about $4\pi \times 10^{-14}$ cm.

For the electron of spin ½ the Dirac theory shows that the electron should have a certain magnetic moment. It turns out that with such a magnetic moment the negative magnetic energy almost perfectly cancels the positive electric energy. The difference still diverges, although only logarithmically. If a cutoff is applied to integrals over wavelengths, at the wavelength limit suggested above, the correction to the electron mass is only about 3 per cent, but there is no way to test this, for we do not recognize a neutral counterpart to the electron.

For the proton the anomalous magnetic moment is so high that the magnetic energy exceeds the electric energy and the correction can be negative. The neutron is a magnet, too, so its correction is also negative. Since the proton moment is higher, the fact that the neutron is heavier than the proton might be explained. If the integrals are cut off at an energy of the order of a proton mass, the difference comes out correctly, but this is too crude a way to calculate such an accurately known energy as the 782.71 ± 0.40 kev† equivalent to this mass difference. These mass differences (of proton and neutron, of charged and neutral pion, of positive, neutral, and negative sigma mesons, and of charged and neutral K mesons, etc.) present a serious challenge to modern physics and possibly point to the failure of quantum electrodynamics to give us a complete theory for calculating electromagnetic effects. We do not know whether it is truly quantum

† Page 354 of Everling et al., *Nucl. Phys.*, vol. 15, p. 342, 1960.

electrodynamics or our assumptions about the distribution of charge inside the particles which are at fault. Only when we have a more complete theory of these particles and their interactions will we be able to determine the limitations, if any, of our present theory of quantum electrodynamics.

9-6 THE LAMB SHIFT

According to the Schrödinger equation, the second level of the hydrogen atom is degenerate. The $2s$ and $2p$ levels occur at the same energy. Likewise, for the Dirac equation there is a degenerate pair $2s_{1/2}$ and $2p_{1/2}$. But Lamb and Rutherford found in 1946 that there is indeed a small separation (about 1 part in 3×10^6) with the $2s_{1/2}$ lying higher by a frequency of 1,057.1 megacycles.

Although theorists reasoned that such an energy difference might arise from effects of the term I, the infinities of the divergent integrals confused all attempts to calculate the difference until the work of Bethe and Weisskopf in 1947. They reasoned as follows:

First, since

$$\frac{1}{E_M - E_N - \hbar kc} = \frac{1}{\hbar kc} \frac{E_M - E_N}{E_M - E_N - \hbar kc} - \frac{1}{\hbar kc} \tag{9-72}$$

the energy (9-71) was expressed as the sum of three terms,

$$\delta E = \delta E' + \delta E'' + \delta E''' \tag{9-73}$$

where

$$\delta E' = \frac{2\pi e^2}{m^2c^2} \int \frac{d^3k}{(2\pi)^3k^2} \sum_N \frac{(E_M - E_N)(|\mathbf{p}_1 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM} + |\mathbf{p}_2 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM})}{E_M - E_N - \hbar kc} \tag{9-74}$$

$$\delta E'' = -\frac{2\pi e^2}{m^2c^2} \int \frac{d^3k}{(2\pi)^3k^2} \sum_N (|\mathbf{p}_1 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM} + |\mathbf{p}_2 e^{-i\mathbf{k}\cdot\mathbf{R}}|^2_{NM}) \tag{9-75}$$

$$\delta E''' = \frac{2\pi e^2\hbar}{mc} \int \frac{d^3k}{(2\pi)^3k} \tag{9-76}$$

The term $\delta E'''$ and the infinity from the Coulomb term δE_c are independent of the state of the electron. They would (it was hoped) be made finite in some future theory. It would contribute some δm to the rest mass of an electron. If m_0 is the mechanical mass, the true experimental mass $m = m_0 + \delta m$, where $\delta mc^2 = \delta E''' + \delta E_c$. In the total energy (including the rest energy of the particles and the binding energy) of the hydrogen atom, such a rest-energy correction to the energy is, of course, expected, but we have already included it when we measure all binding energies relative to the free-particle ionized state. The δm term is thus identified, because it is the only term for an electron at rest, and it is independent of the motion or state of the electron.

The term $\delta E''$ could be simplified; for the sum on N could be taken to give $(p_1^2 + p_2^2)_{MM}$ (by the law of matrix multiplication). When $\mathbf{k}$ is integrated over all directions, this becomes $\frac{2}{3}(\mathbf{p} \cdot \mathbf{p})_{MM}$, and

$$\delta E'' = -\frac{(\mathbf{p} \cdot \mathbf{p})_{MM}}{2m} \frac{8\pi e^2}{3mc^2} \int \frac{d^3k}{(2\pi)^3 k^2} \tag{9-77}$$

Again it was hoped that some day this term would be finite. It exists even for a free electron. It is interpreted as follows: The mechanical kinetic energy $p^2/2m_0$ would be altered (if the mass is altered) to the expression

$$\frac{p^2}{2m} \simeq \frac{p^2}{2m_0}\left(1 - \frac{\delta m}{m_0}\right) \tag{9-78}$$

and the term $\delta E'''$ must represent $-(p^2/2m_0)\ \delta m$. But we have already taken this term into account; for we calculate the Schrödinger energy levels with $p^2/2m$, where m is the experimental mass. The term is identified because it is the only extra term for a moving free electron and it is proportional to the kinetic energy.[1] Finally, even though these terms may be interpreted wrongly, when we calculate the difference of the values of δE for the 2s and 2p states, the terms will drop out, because $\delta E'''$ and δE_c are the same for all states and $\delta E''$ is also the same, since $(p^2/2m)_{MM}$ turns out to be the same for the two states 2s and 2p.

In the final term $\delta E'$ the argument was made that the dipole approximation would suffice. Then the matrix elements are independent of $\mathbf{k}$, and since

$$\int \frac{d^3\mathbf{k}}{k^2} \frac{1}{E_M - E_N - \hbar kc} = \frac{4\pi}{\hbar c} \ln \frac{\hbar k_{\max} c}{E_M - E_N} \tag{9-79}$$

we get

$$\delta E' = \frac{e^2}{\pi m^2 \hbar c^3} \sum_M \left[\ln\left(\frac{\hbar k_{\max} c}{E_M - E_N}\right)\right] (E_M - E_N) \tfrac{2}{3} |\mathbf{p}_{NM}|^2 \tag{9-80}$$

---

[1] The δm implied by Eq. (9-77) is $(8\pi e^2/3c^2) \int d^3k/k^2$ and is *not* equal to the δm obtained from the $\delta E/c^2$ for a static electron. This is because we are limited to a nonrelativistic approximation. When a fully relativistic analysis is carried out, the two ways of calculating δm agree.

Since the states and the matrix elements are known for hydrogen, the sum can be worked out. The only question is the value of $\hbar k_{\max}c$. Bethe argued that the nonrelativistic approximation is at fault here and that, if the full relativistic calculation were made, $\hbar k_{\max}c$ would turn out to be of the order mc^2. Putting $\hbar k_{\max}c = mc^2$ gave about 1,000 megacycles, so Bethe knew he was on the right track.

The remaining problem was to make a relativistic calculation with the Dirac wave function and states. Only in this way could a precise determination of the effective $k_{\max}$ be made. This turned out to be quite confusing, for it was hard to identify the various infinite terms. It would not do to simply cut them off at some maximum momentum and take the difference; for it is not necessarily a relativistically invariant procedure, because it deals with momentum and energy in different ways. (One consequence of this has already been pointed out in the footnote.) One method for resolving the confusion was developed by Schwinger, who showed how the relativistic symmetry could be kept clear throughout the calculation and the infinite terms identified. Another method worked out by Feynman was to give a relativistically invariant procedure to cut off the infinite integrals. Here we shall illustrate the latter method.

The total effect of the electromagnetic field, which this time includes the Coulomb interaction, is represented by an extra term $I + S_c$ in the action. The relativistic invariance of an expression for I like Eq. (9-64) will not be self-evident, since that formula is expressed in terms of $\mathbf{k}$ and t instead of either R and t or $\mathbf{k}$ and ω. Let us represent I in terms of frequency ω and wave number $\mathbf{k}$ variables. First note that

$$\int_{-\infty}^{\infty} e^{-ikc|\tau|}e^{-i\omega\tau}\,d\tau = \frac{2ikc}{\omega^2 - k^2c^2 + i\epsilon} \tag{9-81}$$

or

$$e^{-ik|t-s|c} = \int \frac{2ikc\,d\omega/2\pi}{\omega^2 - k^2c^2 + i\epsilon}\,e^{i\omega(t-s)} \tag{9-82}$$

Suppose we define

$$\begin{aligned} j(\mathbf{k},\omega) &= \int j_{\mathbf{k}}(t)e^{+i\omega t}\,dt \\ &= \int\int j(\mathrm{R},t)e^{-(\mathbf{k}\cdot\mathrm{R}-\omega t)}\,d^3\mathrm{R}\,dt \end{aligned} \tag{9-83}$$

Then I becomes

$$I = -\,2\pi \int \frac{|j_1(\mathbf{k},\omega)|^2 + |j_2(\mathbf{k},\omega)|^2}{\omega^2 - k^2c^2 + i\epsilon}\,\frac{d^3\mathbf{k}\,d\omega}{(2\pi)^4} \tag{9-84}$$

The relativistic symmetry of this expression in ω and $\mathbf{k}$ is already clear; for $\omega^2 - \mathbf{k}^2c^2$ is invariant to the Lorentz transformation. The currents, however, do not appear in a relativistically symmetrical manner. We would have expected an invariant combination like $c^2\rho^2 - \mathbf{j}\cdot\mathbf{j}$, since ρc and $\mathbf{j}$ form a four-vector. But if we define

$$\begin{aligned}\rho(\mathbf{k},\omega) &= \int \rho_{\mathbf{k}}(t)e^{+i\omega t}\,dt \\ &= \int\int \rho(\mathbf{R},t)e^{-i(\mathbf{k}\cdot\mathbf{R}-\omega t)}\,d^3\mathbf{R}\,dt \end{aligned} \tag{9-85}$$

the Coulomb portion of the action is

$$S_c = \int \frac{[\rho(\mathbf{k},\omega)]^2}{k^2}\,d\omega = \int \frac{(\omega\rho/k)^2 - \rho^2c^2}{\omega^2 - k^2c^2}\,d\omega \tag{9-86}$$

the last resulting simply by multiplying numerator and denominator by $\omega^2/k^2 - c^2$. But the law of conservation of current

$$-\frac{\partial \rho}{\partial t} = \nabla\cdot\mathbf{j} \tag{9-87}$$

becomes

$$\omega\rho(k,\omega) = \mathbf{k}\cdot\mathbf{j}(\mathbf{k},\omega) \tag{9-88}$$

Alternatively, if we call j_3 the component of $\mathbf{j}$ in the direction of $\mathbf{k}$, $\omega\rho/k = j_3$ and we have altogether

$$I + S_c = -2\pi \int \frac{|j_1(\mathbf{k},\omega)|^2 + |j_2(\mathbf{k},\omega)|^2 + |j_3(\mathbf{k},\omega)|^2 - c^2|\rho(\mathbf{k},\omega)|^2}{\omega^2 - k^2c^2 + i\epsilon} \times \frac{d^3\mathbf{k}\,d\omega}{(2\pi)^4} \tag{9-89}$$

The sum of the three j terms is just $\mathbf{j}(\mathbf{k},\omega)\cdot\mathbf{j}(\mathbf{k},\omega)$, and the four-dimensional invariance is evident.

A suggestion is now made that in view of our present ignorance, convergence of the integrals can be made artificially by supplying an additional factor such as

$$\left(\frac{-\Lambda^2}{\omega^2 - k^2c^2 - \Lambda^2 + i\epsilon}\right)^2$$

in the integrand of Eq. (9-89), where Λ is some very high frequency. For small values of ω and k this factor is unity, whereas for high values it cuts off the integral. Furthermore, such a factor clearly does not destroy the relativistic invariance of the expression. All physical quantities are to be calculated by assuming $I + S_c$ contains this cutoff factor. If they are insensitive to Λ for large Λ (like the Lamb shift), the theoretical value is to be trusted. If, on the other hand, the result depends sensitively on Λ (as does the charged and neutral pion mass difference) no quantitative meaning can be given to the result, for the cutoff function is arbitrary and is not completely satisfactory. This is the present state of quantum electrodynamics.

Problem 9-11 Show that the cutoff function is not completely satisfactory by arguing that γ calculated in the manner of Sec. 9-4 would be altered by the cutoff, yet the probability of emission of a real photon would not be so altered because for such a photon $\omega = kc$ and the modifying cutoff factor is exactly 1. Thus the balance of probability would not result (i.e., the probability that the atom emits plus the probability that it does not emit would no longer add to unity). The difficulty suggested by this problem has never been solved. No modification of quantum electrodynamics at high frequencies is known which simultaneously makes all results finite, maintains relativistic invariance, and keeps the sum of probabilities over all alternatives equal to unity.

Problem 9-12 Transform $I + S_c$ into space coordinates by using

$$\int e^{i(\mathbf{k}\cdot\mathbf{R}-\omega t)} \frac{c\, d^3\mathbf{k}\, d\omega/(2\pi)^4}{(2\pi)^4(\omega^2 - k^2c^2 + i\epsilon)} = \frac{i}{(t^2c^2 - R^2 - i\epsilon)(2\pi)^2}$$

$$= \frac{1}{4\pi}\,\delta_+(t^2c^2 - R^2) \qquad (9\text{-}90)$$

[*Note:* The function $-\pi i/(x - i\epsilon)$ is often written as $\delta_+(x)$, and we have introduced that convention here.] Then find

$$I + S_c = \frac{1}{2c}\int [c^2\rho(\mathbf{R}_1,t_1)\rho(\mathbf{R}_2,t_2) - \mathbf{j}(\mathbf{R}_1,t_1)\cdot\mathbf{j}(\mathbf{R}_2,t_2)]$$

$$\times\, \delta_+[(t_1 - t_2)^2c^2 - |\mathbf{R}_1 - \mathbf{R}_2|^2]\, d^3\mathbf{R}_1\, d^3\mathbf{R}_2\, dt_1\, dt_2 \qquad (9\text{-}91)$$

9-7 THE EMISSION OF LIGHT

In Sec. 9-4 we found an expression for the amplitude that the matter system would do something when interacting with an electromagnetic field, as shown in Eq. (9-90) and the following development. This derivation was restricted to the special case that the field is both initially and finally in the vacuum state with no photons present. The result was that the action S_{mat} in the path integrals must be replaced by an effective action $S'_{mat} = S_{mat} + I$. In a more general case, photons were present, both initially and finally. As an illustration suppose that initially no photons are present but in the final state there is just one photon of momentum $\hbar\mathbf{L}$ and polarization 1. The only change which this makes in our previous calculation is the change in the integral defining S, that is, Eq. (9-61). We shall now use

$$X' = \frac{\pi}{k} \int e^{(i/\hbar)(S_{mat}+S_{rad})} \, \mathcal{D}a_{1k} \, \mathcal{D}a_{2k} \tag{9-92}$$

where the path integral is carried out between a vacuum initial state and final state consisting of a vacuum plus 1 photon. Then every oscillator, except 1L, goes from the initial state $n = 0$ to the final state $n = 0$, so the factor $X_{1\mathbf{k}}$ for all of these oscillators is unchanged. Only the contribution from the single oscillator 1L is altered; for it now becomes

$$X'_{1\mathbf{k}} = \int \left(\exp \left\{ \frac{i}{\hbar} \int \left[\sqrt{4\pi} \, (j^*_{1\mathbf{L}} a_{1\mathbf{L}} + j_{1\mathbf{L}} a^*_{1\mathbf{L}}) + \dot{a}^*_{1\mathbf{L}} \dot{a}_{1\mathbf{L}} - k^2 c^2 a^*_{1\mathbf{L}} a_{1\mathbf{L}} - \frac{\hbar L c}{2} \right] dt \right\} \right) \mathcal{D}a_{1\mathbf{L}} \tag{9-93}$$

This expression is the same as Eq. (9-63) except that the oscillator path is taken between the state $n = 0$ and the state $n = 1$ instead of $n = 0$ to $n = 0$ as in the previous expression. We worked out the behavior of a forced harmonic oscillator in Sec. 8-9, and we can use the results of that section to write

$$X'_{1\mathbf{L}} = \left(\sqrt{\frac{2\pi\hbar}{Lc}} \int j_{1\mathbf{L}} e^{iLct} \, dt \right) X_{1\mathbf{L}} \tag{9-94}$$

where $X_{1\mathbf{L}}$ is the $n = 0$ to $n = 0$ factor previously calculated. Therefore, evidently the complete factor X' is simply the original factor X multiplied by

$$\sqrt{\frac{2\pi\hbar}{Lc}} \int j_{1\mathbf{L}} e^{iLc} \, dt$$

and we find for the amplitude

$$\text{Amplitude} = \sqrt{\frac{2\pi\hbar}{Lc}} \int e^{(i/\hbar)(S_{mat}+I)} \int j_{1\mathbf{L}} e^{iLct} \, dt \, \mathcal{D}q \tag{9-95}$$

The perturbation theory expression which we previously evaluated was equivalent to the transition element

$$\sqrt{\frac{2\pi\hbar}{Lc}} \int e^{(i/\hbar)S_{mat}} \int j_{1L}e^{iLct}\,dt\,\mathfrak{D}q \tag{9-96}$$

so we see that the net result is the same as that given by the perturbation theory with the following exception. The transition amplitude must be calculated with the complete effective action $S'_{mat} = S_{mat} + I$ instead of just S_{mat}. The effect of I is to change the energy levels a bit as we discussed, but also to make the energy values effectively complex. The result is that the emitted light gives a spectral line which appears to have a little width, which is called the natural width of the line. We shall not go any further into the details of this calculation but leave the subject and the generalizations to a number of photons both entering and leaving the system to those who wish to study quantum electrodynamics specifically in more detail.

9-8 SUMMARY

Review of the Approach. In this chapter we have a considerable amount of analysis of the quantum electromagnetic field. It is worthwhile looking back again to see the central ideas and results. The separation of the Coulomb interaction and the use of running waves are technical ways of accomplishing our ends, but the essential result is the formula of Eq. (9-89) [or its equivalent, Eq. (9-91)]. Let us review this result from the more general point of view exemplified by the ideas of Eq. (9-1).

Suppose we have a system which can be described by an action

$$S = S_1(\mathbf{q}) + S_2(\mathbf{q},\mathbf{A},\phi) + S_3(\mathbf{A},\phi) \tag{9-97}$$

where $S_1(\mathbf{q})$ is the action of the matter alone, S_2 is the interaction of matter and field, and S_3 is the action of the field alone, and where $\mathbf{q}$ stands for all the coordinates of the matter, with $\mathbf{A}$, ϕ describing the field. Then the amplitude for any event results from evaluating an integral like

$$K = \int \exp\left\{\frac{i}{\hbar}[S_1(\mathbf{q}) + S_2(\mathbf{q},\mathbf{A},\phi) + S_3(\mathbf{A},\phi)]\right\} \mathfrak{D}\mathbf{q}\,\mathfrak{D}\mathbf{A}\,\mathfrak{D}\phi \tag{9-98}$$

subject to the boundary conditions of the problem in question. In this summary we shall assume that conditions for the field are that initially and finally no photons are present (i.e., ground state to ground state for the field), and we abbreviate this set of conditions as *gnd-gnd*.

Later on we shall consider the consequences of integrating q first and A, ϕ last. What we have done so far corresponds instead to integrating A, ϕ first and reserving the q integral as a subsequent step.

Usually $S_2(\mathbf{q},\mathbf{A},\phi)$ is linear in the field variables A, ϕ and can be written as

$$S_2 = \int [\rho(\mathbf{R},t)\phi(\mathbf{R},t) - \mathbf{j}(\mathbf{R},t) \cdot \mathbf{A}(\mathbf{R},t)]\, d^3\mathbf{R}\, dt \tag{9-99}$$

where ρ and j are the electric charge and current density, which depend on the q only. The integral on A, ϕ is then easily performed because it is a gaussian integral. It is the burden of Eq. (9-91) to tell us the value of this integral, namely,

$$\int_{gnd}^{gnd} \left(\exp\left\{\frac{i}{\hbar}\left[S_3(\mathbf{A},\phi) + \int (\rho\phi - \mathbf{j}\cdot\mathbf{A})\, d^3\mathbf{R}\, dt\right]\right\}\right) \mathfrak{D}\mathbf{A}\, \mathfrak{D}\phi = e^{(i/\hbar)J} \tag{9-100}$$

where J, which we called $I + S_c$ in Eq. (9-91), is

$$J = i\int [c^2\rho(\mathbf{R}_1,t_1)\rho(\mathbf{R}_2,t_2) - \mathbf{j}(\mathbf{R}_1,t_1)\cdot\mathbf{j}(\mathbf{R}_2,t_2)]\delta_+[(t_1 - t_2)^2c^2 - |\mathbf{R}_1 - \mathbf{R}_2|^2]\, d^3\mathbf{R}_1\, d^3\mathbf{R}_2\, dt_1\, dt_2 \tag{9-101}$$

for *any* functions ρ, j of R, t. The expression for Eq. (9-101) in momentum space appears in Eq. (9-89).

In the applications of Eq. (9-98) these ρ, j are some function of q and $\dot{\mathbf{q}}$, so we obtain the result that

$$K(gnd,gnd) = \int \left(\exp\left\{\frac{i}{\hbar}[S_1(\mathbf{q}) + J(\mathbf{q})]\right\}\right) \mathfrak{D}\mathbf{q} \tag{9-102}$$

where the functional of q, $J(\mathbf{q})$ is given by Eq. (9-101) with the correct ρ, j substituted. This summarizes the results for *gnd-gnd* transitions. We express the modifying effect of the field on the action of the particles by the addition of $J(\mathbf{q})$ to $S_1(\mathbf{q})$. The central formula for electrodynamics then is the general result of Eqs. (9-100) and (9-101).

General Formulation of Quantum Electrodynamics. It is also of interest to pursue these matters in a different direction, by integrating over the matter coordinates first, and leave the field variables for later. We shall limit ourselves to a brief general description of what results from this procedure. If in Eq. (9-98) we contemplate integrating q first, the factor $e^{(i/\hbar)S_3}$ is a constant and can be left out. We can therefore write Eq. (9-98) this way: If we define

$$T[\mathbf{A},\phi] = \int \left(\exp\left\{\frac{i}{\hbar}[S_1(\mathbf{q}) + S_2(\mathbf{q},\mathbf{A},\phi)]\right\}\right) \mathfrak{D}\mathbf{q} \tag{9-103}$$

then

$$K = \int \left\{ \exp\left[\frac{i}{\hbar} S_3(\mathbf{A},\phi)\right]\right\} T[\mathbf{A},\phi]\, \mathfrak{D}\mathbf{A}\, \mathfrak{D}\phi \tag{9-104}$$

This K gives us the amplitude that the particle goes through a certain motion *and* the field undergoes a certain transition. Like all other amplitudes, it is the sum over all possible alternatives. Each separate alternative is constructed as the amplitude $T[\mathbf{A},\phi]$ for the motion in a particular field $\mathbf{A}$, ϕ times the amplitude $e^{iS_3/\hbar}$ that the field is $\mathbf{A}$, ϕ. In carrying out the sum, we sum over all possible fields $\mathbf{A}$, ϕ.

This law, given by Eq. (9-104), is the general fundamental rule for all of quantum electrodynamics. It is a correct formulation even when the functional $T[\mathbf{A},\phi]$, the amplitude for the motion of the particles in an external potential $\mathbf{A}$, ϕ, cannot be represented as a path integral. For example, for a relativistic particle with spin (described by the Dirac equation), the quantity $T[\mathbf{A},\phi]$ cannot be described by a simple path integral based on any reasonable action. However, it is possible to calculate $T[\mathbf{A},\phi]$ by other means, for example, from the Dirac equation. After the form of this functional has been derived, the amplitude K can be worked out, in principle, from Eq. (9-104).

In stating the law of quantum electrodynamics in the form of Eq. (9-104), we have isolated the behavior of the electromagnetic field from the behavior of the particle (or system of particles) on which it acts. That such an isolation can be carried out is an important result. For example, the functional $T[\mathbf{A},\phi]$ may represent the behavior of a nucleus whose properties are not completely known. However, if we know only the behavior of the nucleus in an external field, then we can solve quantum-electrodynamic problems involving nuclei.

Of course, to use Eq. (9-104) strictly, T must be known as a functional of $\mathbf{A}$, ϕ for all $\mathbf{A}$, ϕ, but this much information is rarely available. Even if it is available, the path integral over $\mathbf{A}$, ϕ may not be easy. But in practice the formula is very useful. Sometimes T can be approximated by an exponential, linear in $\mathbf{A}$, ϕ, of exactly the form of Eq. (9-99). Then the result is obtained directly from the general formula of Eqs. (9-100) and (9-101). More often, T can be represented by a sum, or integral, over such exponential forms with various ρ, $\mathbf{j}$ and the result of Eq. (9-104) is a corresponding sum or integral over expressions containing $e^{(i/\hbar)J}$ with J in Eq. (9-101) involving the corresponding ρ, $\mathbf{j}$.

In most practical situations T can be expressed as a power series in $\mathbf{A}$, ϕ. The first few terms can be found from the theory of the matter considering $\mathbf{A}$, ϕ as a small perturbation. Subsequent substitution into Eq. (9-104) and integration on **A**, ϕ gives a corresponding perturbation expansion (in powers of $e^2/\hbar c$) for K. The necessary path integrals such as

$$\int A_i(\mathbf{R}_1,t_1)\,A_j(\mathbf{R}_2,t_2)\left\{\exp\left[\frac{i}{\hbar}S_3(\mathbf{A},\phi)\right]\right\}\mathfrak{D}\mathbf{A}\,\mathfrak{D}\phi = 2\hbar\delta_+[(t_1-t_2)^2c^2-|\mathbf{R}_1-\mathbf{R}_2|^2]$$

can be discovered by expanding the general formula of Eqs. (9-100) and (9-101) on both sides in powers of ρ, **j** and comparing corresponding terms. We shall not go further into these matters here, but refer the reader to the literature (e.g., sec. 8 of R. P. Feynman, Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction, *Phys. Rev.*, vol. 80, p. 440, 1950).

10

Statistical Mechanics

In preceding chapters we have discussed transitions in which a system goes from one known state to another. In most physically realistic situations the initial state is not completely determined. The system may be in one or another state with different probabilities associated with each. In this case the final state is equally uncertain, being that set of states resulting from the various possible initial states with the corresponding probabilities. Or we may not be interested in the probability to go to just one special final state, but rather the chance to end up in any one of a set of such states.

An especially interesting case of statistical uncertainty of states is that corresponding to thermal equilibrium at some temperature T. A quantum-mechanical system in thermal equilibrium can exist in one or another energy level. The results of quantum-statistical mechanics show that the probability that a system is in a state of energy E is proportional to $e^{-E/kT}$, where kT measures the temperature in natural energy units. (The conversion factor k, known as Boltzmann's constant, is 1.38047×10^{-16} erg/°C, or 1 ev per 11,606°K.)

In this book we shall neither derive nor discuss this exponential distribution law. We emphasize that the energy E is the energy of the entire system. If an energy level is degenerate, then each state at that particular level has equal probability. This means that the total probability for the system to have the particular energy *value* is enhanced by a factor corresponding to the number of states in the degenerate level.

The exponential law given above is not yet a true probability distribution, since it has not been normalized. The normalizing factor can be written $1/Z$, so that the probability that a system should be in the state of energy E_i (assumed nondegenerate for the time being) is

$$p_i = \frac{1}{Z} e^{-E_i\beta} \tag{10-1}$$

where $\beta = 1/kT$. This means

$$Z = \sum_i e^{-E_i\beta} \tag{10-2}$$

An equivalent normalization consists in defining an energy F such that

$$p_i = e^{-\beta(E_i-F)} \tag{10-3}$$

F is called the *Helmholtz free energy*. Its value is, of course, dependent on the temperature T, although the various energy levels E_i do not depend on T. It is evident that

$$Z = e^{-\beta F} \tag{10-4}$$

10-1 THE PARTITION FUNCTION

The physical properties of a system in thermal equilibrium can be derived from the exponential distribution function. Suppose A is the measure of some property and that its mean value in the ith state is

$$A_i = \int \phi_i^* A \phi_i \, dV \tag{10-5}$$

where the integral is taken over the volume V of the system. Then the statistical average for A for the whole system is

$$\bar{A} = \sum p_i A_i = \frac{1}{Z} \sum A_i e^{-E_i\beta} \tag{10-6}$$

For example, the average or expected value of the energy itself is

$$\begin{aligned} U &= \sum p_i E_i = \frac{1}{Z} \sum E_i e^{-\beta E_i} \\ &= \sum E_i e^{-\beta(E_i - F)} \end{aligned} \tag{10-7}$$

If the normalizing factor Z is known as a function of the temperature, the sum of Eq. (10-7) can be easily evaluated. From Eq. (10-2) we have

$$\sum E_i e^{-\beta E_i} = -\frac{\partial Z}{\partial \beta} = kT^2 \frac{\partial Z}{\partial T} \tag{10-8}$$

This means that

$$\begin{aligned} U &= \frac{kT^2}{Z} \frac{\partial Z}{\partial T} = kT^2 \frac{\partial \ln Z}{\partial T} = F - T \frac{\partial F}{\partial T} \\ &= \frac{\partial}{\partial \beta} (\beta F) \end{aligned} \tag{10-9}$$

We have written the derivatives with respect to the temperature as partial derivatives because other variables, such as the volume of the system or any external fields, which determine the energy levels are all held fixed.

It is interesting to see what happens to the expected value of the energy if some other variable such as the volume is changed. Suppose the system is in a particular state E_i and we make a small change in the value of a certain parameter, say, α. Using a first-order perturbation principle, we find that the first-order change in energy is equal to the expected value of the first-order change in the hamiltonian. That is,

$$E_i + \Delta E_i = \int \phi_i^*(H + \Delta H)\phi_i \, dV$$
$$\Delta E_i = \int \phi_i^* \, \Delta H \, \phi_i \, dV \tag{10-10}$$

Using the language of classical physics, we would say that the ratio $\Delta H/\Delta\alpha$ is the "force" associated with the parameter α. In case this parameter is the volume, the force is minus the pressure. That is, we define the concept of force by the relation

Force × change in parameter = change in energy

or

$$f_\alpha = \frac{\partial H}{\partial \alpha} \tag{10-11}$$

As an example, then, if P = pressure and V = volume,

$$-P\,\Delta V = \Delta E \tag{10-12}$$

We write the expected value of the force as

$$\begin{aligned}\bar{f}_\alpha &= \overline{\left(\frac{\partial H}{\partial \alpha}\right)} = \sum p_i \left(\frac{\partial H}{\partial \alpha}\right)_i \\ &= \sum p_i \frac{\partial E_i}{\partial \alpha} = \sum \frac{1}{Z}\frac{\partial E_i}{\partial \alpha} e^{-E_i/kT} = -\frac{kT}{Z}\frac{\partial}{\partial \alpha}\left(\sum e^{-E_i/kT}\right) = -\frac{kT}{Z}\frac{\partial Z}{\partial \alpha}\end{aligned} \tag{10-13}$$

so that

$$\bar{f}_\alpha = -\frac{1}{\beta}\frac{\partial \ln Z}{\partial \alpha} \tag{10-14}$$

where β and other parameters are held constant. Using Eq. (10-4), we can write this as

$$\bar{f}_\alpha = \frac{\partial F}{\partial \alpha} \tag{10-15}$$

If the parameter α is the volume V so that $-\bar{f}_\alpha$ is the pressure P, we have

$$P = -\frac{\partial F}{\partial V} \tag{10-16}$$

When the volume is changed by an infinitesimal amount for a system at a constant temperature, two things happen simultaneously. First, each of the energy levels shifts slightly. Second, for the system to stay in equilibrium at a constant temperature T (maintained by a bath, for example) the probabilities associated with each energy level will change slightly (because the energy of the levels changes). If the only effect were a change in the energy of the levels, then the change in the total energy of the system would correspond to this change averaged over all the levels. From our foregoing discussion this corresponds to pressure times the change in volume. However, to keep the temperature fixed, some readjustment of the population of states must occur. Thus the total energy must make an additional change which we shall call dQ. This additional energy is given up to the external system (the bath) which maintains the temperature, and it is called the heat exchanged. Thus

$$dU = -P\,dV + dQ \tag{10-17}$$

We can find dQ easily from the expression for U given in Eq. (10-7). When V is altered by the change dV, then each energy level E_i undergoes the change dE_i and the Helmholtz free energy changes by an amount dF. Thus the total energy changes by the amount

$$dU = \Sigma\, dE_i\, e^{-\beta(E_i-F)} + \beta\, dF\, \Sigma E_i e^{-\beta(E_i-F)} - \beta \Sigma E_i\, dE_i\, e^{-\beta(E_i-F)} \tag{10-18}$$

The first term in this expression is the expected value of dE_i, which is $-P\,dV$, as we have already explained. The remaining two terms constitute dQ. These two terms can also be expressed with the derivatives of the sum in Eq. (10-2), and ultimately in terms of F. In fact, we find

$$dQ = T\frac{\partial^2 F}{\partial T\,\partial V}\,dV \tag{10-19}$$

That this is true can be seen also from Eq. (10-17), which gives

$$\begin{aligned} dQ &= \frac{\partial U}{\partial V} + P \\ &= \frac{\partial}{\partial V}\left(F - T\frac{\partial F}{\partial T}\right) - \frac{\partial F}{\partial V} \\ &= -T\frac{\partial^2 F}{\partial T\,\partial V} \end{aligned} \tag{10-20}$$

Equation (10-19) gives the heat exchanged dQ in changing the volume by the amount dV while keeping the temperature constant. If we change any other parameter, we shall arrive at an analogous result. For example, if we change the temperature T while holding the volume V constant, the heat exchanged is equal to the change in total energy. That is,

$$\Delta Q = \frac{\partial U}{\partial T}\,\Delta T = \frac{\partial}{\partial T}\left(F - T\frac{\partial F}{\partial T}\right)\Delta T = -T\frac{\partial^2 F}{\partial T^2}\,\Delta T \tag{10-21}$$

In general, then, we have the result

$$\Delta Q = -T\left(\frac{\partial^2 F}{\partial T\,\partial V}\,\Delta V + \frac{\partial^2 F}{\partial T\,\partial \alpha}\,\Delta\alpha + \frac{\partial^2 F}{\partial T^2}\,\Delta T\right) \tag{10-22}$$

The right-hand side of Eq. (10-22) is of the form T times the total change in a quantity $\mathsf{S} = -(\partial F/\partial T)$, which is called the *entropy*. That is,

$$\Delta Q = T\,\Delta\mathsf{S} \tag{10-23}$$

$$\mathsf{S} = -\frac{\partial F}{\partial T} \tag{10-24}$$

$$U = F - T\mathsf{S} \tag{10-25}$$

It is evident that the standard thermodynamic quantities, internal energy, entropy, pressure, etc., can all be evaluated if a single function, the partition function Z, is known in terms of the temperature, volume, external field, etc. The thermodynamic quantities are obtained simply by differentiating Z, or equivalently, the free energy F.

The determination of some physical quantities, even for a system in thermal equilibrium, requires more information than only the partition function. For example, suppose the system is in a configuration space with a coordinate x and we ask: What is the probability of finding the system at location x? We know that if the system is in the single state defined by the wave function $\phi_i(x)$, the probability of observing x is the absolute square of the wave function $\phi_i^*(x)\phi_i(x)$. Thus, averaging over all possible states, the probability of observing x is

$$P(x) = \frac{1}{Z}\sum_i \phi_i^*(x)\phi_i(x)e^{-\beta E_i} \tag{10-26}$$

In the general case, if we are interested in any quantity A, then the expected value is given by

$$\bar{A} = \frac{1}{Z}\sum_i A_i e^{-\beta E_i} = \frac{1}{Z}\sum_i \int \phi_i^*(x)\,A\,\phi_i(x)e^{-\beta E_i}\,dx \tag{10-27}$$

It is evident that the expected values of all such quantities could be obtained if we knew the function

$$\rho(x',x) = \sum_i \phi_i(x')\phi_i^*(x)e^{-\beta E_i} \tag{10-28}$$

This suffices since the function A appearing in the integral of Eq. (10-27) is an operator which operates only on the ϕ_i of that expression, and not on ϕ_i^*. Using the quantity $\rho(x',x)$, we can imagine A to act on x' only, after which we set x' equal to x in the form $A\rho(x'x)$, and finally integrate over all values of x. This process is called finding the *trace* of $A\rho$.

From the definition of $\rho(x',x)$ it is clear that

$$P(x) = \frac{1}{Z}\rho(x,x) \tag{10-29}$$

and since the probability $P(x)$ is normalized, so that the integral over all x gives 1, we have

$$Z = \int \rho(x,x)\,dx = \mathrm{Tr}\,[\rho] \tag{10-30}$$

where Tr is an abbreviation for the word "trace." The quantity $\rho(x',x)$ is called the *density matrix* [more precisely, the statistical density matrix for temperature T; the term "density matrix" also has a wider use for general systems in or out of thermal equilibrium and is usually used for the normalized version of the function $\rho(x',x)$, that is, the function we would write as $\rho(x',x)/Z$]. The general problem of statistical mechanics is to evaluate Eq. (10-28) to find the density matrix. If we are interested only in conventional thermodynamic variables, we need only the trace or diagonal sum of the density matrix, which gives us the partition function Z.

10-2 THE PATH INTEGRAL EVALUATION

The formulation of the density matrix given in Eq. (10-28) bears a close resemblance to the general expression for the kernel, which was derived in Chap. 4 and given in Eq. (4-59) as

$$K(x_2,t_2;x_1,t_1) = \sum_j \phi_j(x_2)\phi_j^*(x_1)e^{-(i/\hbar)E_j(t_2-t_1)} \tag{10-31}$$

The validity of this expression is restricted to situations in which the hamiltonian is constant in time and $t_2 > t_1$. However, this situation is implied in statistical mechanics; for only if the hamiltonian is constant in time can equilibrium be achieved. The difference between the form of Eq. (10-31) and that of Eq. (10-28) is in the argument of the exponential. If the time difference $t_2 - t_1$ of Eq. (10-31) is replaced by $-i\beta\hbar$, we see that the expression for the density matrix is formally identical to the expression for the kernel corresponding to an imaginary negative time interval.

We can develop the similarity between these two expressions from another point of view. Suppose we write the density matrix in a way which makes it look a little bit more like a kernel, thus, $k(x_2,u_2;x_1,u_1)$ for $\rho(x_2,x_1)$, where

$$k(x_2,u_2;x_1,u_1) = \sum_i \phi_i(x_2)\phi_i^*(x_1)e^{-[(u_2-u_1)/\hbar]E_i} \tag{10-32}$$

Then if $x_2 = x'$, $x_1 = x$, $u_2 = \hbar\beta$, and $u_1 = 0$, Eq. (10-32) becomes identical with Eq. (10-28).

If we differentiate k partially with respect to u_2, we get

$$-\hbar\frac{\partial k}{\partial u_2} = \sum_i E_i\phi_i(x_2)\phi_i^*(x_1)e^{-[(u_2-u_1)/\hbar]E_i} \tag{10-33}$$

But now we recall that $E_i\phi_i(x') = H\phi_i(x')$; so if we understand H_2 to imply operations only upon the variables x_2, we can write

$$-\hbar\frac{\partial k(2,1)}{\partial u_2} = H_2 k(2,1) \tag{10-34}$$

or, to put the same thing another way,

$$-\frac{\partial\rho(2,1)}{\partial\beta} = H_2\rho(2,1) \tag{10-35}$$

We notice that this differential equation for ρ is similar to the Schrödinger equation for the kernel K which was developed in Chap. 4 and given in Eq. (4-25). We can rewrite it here as

$$-\frac{\hbar}{i}\frac{\partial K(2,1)}{\partial t_2} = H_2K(2,1) \qquad \text{for } t_2 > t_1 \tag{10-36}$$

We found in Chap. 4 that the kernel $K(2,1)$ is Green's function for Eq. (10-36). In the same sense the density matrix $\rho(2,1)$ is Green's function for Eq. (10-35).

With simple hamiltonians involving only momenta and coordinates we have been able to write the kernel as a path integral. For example, if the hamiltonian is given by

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \tag{10-37}$$

then the solution for the kernel over a very short time interval

$t_2 - t_1 = \epsilon$

is

$$K(2,1) = \sqrt{\frac{m}{2\pi i\hbar\epsilon}} \exp\left[\frac{im}{2\hbar}\frac{(x_2 - x_1)^2}{\epsilon} - \frac{i}{\hbar}\epsilon V\left(\frac{x_2 + x_1}{2}\right)\right] \tag{10-38}$$

which can be directly verified by substitution into Eq. (10-37). By building up a product of many kernels of the form of Eq. (10-38) and taking the limit as the time interval ϵ goes to 0 and the number of terms in the product becomes infinite, we have produced a path integral describing the kernel over a finite period of time. We can produce a solution to Eq. (10-34) in the same manner. The solution for an infinitesimal interval of $u_2 - u_1 = \eta$ is given by substituting $\epsilon = -i\eta$ into Eq. (10-38). Thus

$$k(x_2,\eta;x_1,0) = \sqrt{\frac{m}{2\pi\eta\hbar}} \exp\left\{-\frac{(m/2\eta)(x_2 - x_1)^2 + \eta V[(x_2 + x_1)/2]}{\hbar}\right\} \tag{10-39}$$

That this is a valid solution of Eq. (10-34) can be demonstrated by direct substitution of Eq. (10-39) into Eq. (10-34).

The rule for the combination of functions defined for successive values of u is the same as the rule for the combination of kernels for successive intervals of time. That is,

$$k(2,1) = \int k(2,3)k(3,1)\, dx_3 \tag{10-40}$$

That this result still holds follows from the fact that Eq. (10-33) is a first-order derivative in u. This rule can be used to obtain the path integral to define $k(2,1)$ as

$$k(x_2,u_2;x_1,u_1) = \int \left(\exp\left\{-\sum_{i=0}^{N-1}\left[\frac{m}{2\hbar\eta}(x_{i+1} - x_i)^2 + \frac{\eta}{\hbar}V(x_i)\right]\right\}\right) \times \prod_{i=1}^{N-1}\frac{dx_i}{a} \tag{10-41}$$

The normalizing constant a now becomes

$$a = \sqrt{\frac{2\pi\hbar\eta}{m}} \tag{10-42}$$

and the integral is carried out over all paths going from x_1 to x_2 (that is, x_i is x_1 for $i = 0$ and x_2 for $i = N$) in the interval $u_2 - u_1 = N\eta$.

The result of this derivation is that if we consider a "path" $x(u)$

as a function which gives a coordinate in terms of the parameter u, and if we call $\dot{x}$ the derivative dx/du, then

$$\rho(x_2,x_1) = \int \left(\exp\left\{-\frac{1}{\hbar}\int_0^{\beta\hbar}\left[\frac{m}{2}\,\dot{x}^2(u) + V(x)\right]du\right\}\right)\mathfrak{D}x(u) \qquad (10\text{-}43)$$

This is a very amusing result, because it gives the complete statistical behavior of a quantum-mechanical system as a path integral without the appearance of the ubiquitous i so characteristic of quantum mechanics. (Incidentally, this is not so for a system moving in a magnetic field.) This path integral of Eq. (10-43) is much easier to work with and visualize than the complex integrals which we have studied previously. Here, it is easy to see why some paths contribute very little to the integral; for there are paths for which the exponent is very large and thus the integrand is negligibly small. Furthermore, it is not necessary to think about whether or not nearby paths cancel each other's contributions, since in the present case all contributions add together with some being large and others small.

The parameter u is not the true time in any sense. It is just a parameter in an expression for the density matrix ρ. However, if we wish to make use of analogy in our thinking, we can consider u as the time for a certain path, and in so doing we can state the result given by Eq. (10-43) in a vivid pictorial way. What we are doing is providing a physical analogue for the mathematical expression. We shall call u the "time," leaving the quotation marks to remind us that it is not real time (although u does have the dimensions of time). Likewise $\dot{x}$ will be called the "velocity," $m\dot{x}^2/2$ the "kinetic energy," etc. Then Eq. (10-42) says that the density matrix for a temperature $1/\beta$ is given in the following way:

Consider all the possible paths, or "motions," by which the system can travel between the initial and final configurations in the "time" $\beta\hbar$. The density matrix ρ is a sum of contributions from each motion, the contribution from a particular motion being the "time" integral of the "energy" divided by $\hbar$ for the path in question.

The partition function is derived by considering only those cases in which the final configuration is the same as the initial configuration, and we sum over all possible initial configurations.

Problem 10-1 Show that the expression for the density matrix for a system consisting of a harmonic oscillator is

$$\rho(x',x) = \left(\frac{m\omega}{2\pi\hbar\sinh\omega\beta\hbar}\right)^{1/2}\exp\left\{-\frac{m\omega}{2\hbar(\sinh\omega\beta\hbar)^2}\right.$$
$$\left.\times\left[(x^2+x'^2)\cosh\omega\beta\hbar - 2xx'\right]\right\} \qquad (10\text{-}44)$$

This answer can be compared with the results of Prob. 3-8. Show also that the free energy is $kT \ln [2 \sinh (\hbar\omega/2kT)]$. Check this latter value by a direct evaluation of the sum of Eq. (10-2).

If the temperature is not too low (how low is too low will be discussed later on in this chapter), $\beta\hbar$ is very small. Thus, in calculating the partition function for which $x_1 = x_2$, each path starts from x_1 and in a very short "time" is back to x_1 again. In fact, the paths cannot ever wander very far from x_1, because traveling far away and returning again in the short "time" available requires a high "velocity" and a large "kinetic energy." For such a path the exponential function appearing in Eq. (10-43) becomes negligibly small, and it will contribute a negligible amount to the sum over all paths. Under these circumstances the paths $x(u)$ which must be considered in evaluating $V[x(u)]$ never move very far from the initial point x_1. Thus to a first approximation we can write $V[x(u)] \simeq V[x_1]$ for all paths. In this approximation the potential energy is independent of the path, and the exponential function dependent on the potential can be taken outside the integral. Thus for temperatures which are not too low we have

$$\rho(x_1,x_1) = e^{-\beta V(x_1)} \int_{x_1}^{x_1} \left\{ \exp \left[-\frac{m}{2\hbar} \int_0^{\beta\hbar} \dot{x}^2(u)\, du \right] \right\} \mathfrak{D}x(u) \tag{10-45}$$

In this last expression the path integral is that for a free particle. It can be solved in the same way that we solved the path integral defining the kernel for the motion of a free particle in Chap. 3. The result is

$$\int_{x_1}^{x_2} \left\{ \exp \left[-\frac{m}{2\hbar} \int_0^{\beta\hbar} \dot{x}^2(u)\, du \right] \right\} \mathfrak{D}x(u) = \sqrt{\frac{mkT}{2\pi\hbar^2}} \exp \left[-\frac{mkT(x_2 - x_1)^2}{2\hbar^2} \right] \tag{10-46}$$

If we are interested only in the partition function, we set $x_2 = x_1$ and find

$$\rho(x_1,x_1) = \sqrt{\frac{mkT}{2\pi\hbar^2}}\, e^{-\beta V(x_1)} \tag{10-47}$$

Then, the partition function is the integral of this expression over all possible initial configurations x_1. Thus

$$Z = \sqrt{\frac{mkT}{2\pi\hbar^2}} \int e^{-\beta V(x_1)}\, dx_1 \tag{10-48}$$

This is a formula for the partition function valid in the limit of classical mechanics. It was originally derived, within an uncertain multiplying constant, as a consequence of classical mechanics by Boltzmann. In more complicated cases (e.g., more variables) the classical partition function is simply the product of two factors. The first of these is the path integral which one would get by considering all particles of the system to be free. The second factor is called the configuration integral, and it is just the integral of $e^{-\beta V}$, where V is the potential of the system and depends upon all of the N variables describing the system. For example, for N particles interacting by a potential $V(\mathbf{x}_1,\mathbf{x}_2, \ldots ,\mathbf{x}_N)$, where $\mathbf{x}_a$ is the vector position of particle a, the integral required is

$$\int \{\exp[-\beta V(\mathbf{x}_1,\mathbf{x}_2, \ldots ,\mathbf{x}_N)]\}\, d^3\mathbf{x}_1\, d^3\mathbf{x}_2 \cdots d^3\mathbf{x}_N$$

This simple form for the partition function is only an approximation valid if the particles of the system cannot wander very far from their initial positions in the "time" $\beta\hbar$. The limit on the distance which the particles can wander before the approximation breaks down can be estimated in Eq. (10-46). We see that if the final point differs from the initial point by as much as

$$\Delta x = \frac{\hbar}{\sqrt{mkT}} \tag{10-49}$$

the exponential function of Eq. (10-46) becomes very greatly reduced. From this, we can infer that intermediate points farther than Δx away from the initial and final points can be reached only on paths which do not contribute greatly to the path integral of Eq. (10-43). If the potential $V(x)$ does not alter very much as x moves over this distance, then the classical statistical mechanics is valid.

For a typical solid or liquid at room temperatures, with an atomic mass of about 20, for example, Δx is about 0.1 Å, while the interatomic distances and forces range over one or two angstroms. Thus motions greater than 0.1 Å will not contribute to the density matrix, while the potential function will remain unchanged until the motion of the order of one or two angstroms has been achieved. It is clear that classical statistical mechanics is adequate for such materials.

All of the mysterious transformations between solids, liquids, and gases ordinarily lie in a range where classical statistical mechanics is valid. The mathematical interpretation of all these processes is contained in the problem of evaluating the integral of $e^{-\beta V}$ over the coordinates of all the atoms. That this amazing variety and peculiarity of phenomena comes from just a simple integral is at first surprising, until it is realized that the integral is a multiple integral over a stupendous number of variables. Our usual experience with integrals which involve one or, at most, a few variables of integration does not prepare us for the almost qualitative differences that can arise when the number of variables approaches infinity.

The fascination of the problems in the theory of the solid states, or of liquids and of the condensation of gases, lies, like the behavior of this multiple integral, in the way in which simple descriptions of simple systems when joined together in enormous multiplicity yield such rich phenomena. It is a challenge to the imagination to see how the cooperation between systems can lead to such results. A rough qualitative explanation is readily forthcoming for many of these effects, but the problem of quantitative detail also holds fascination for the theoretical physicist.

There are important statistical phenomena which occur when the classical approximation is not valid. In this case the multiplicity of variables is compounded with the conceptual complexity of quantum mechanics to raise even greater challenges.

Strictly speaking, Eq. (10-48) does contain a little more information than was available to the purely classical statistical mechanics. This is evidenced by the appearance of $\hbar$ in the coefficient in front of the integral. Classical mechanics could not determine the partition function absolutely, but only within an unidentified constant factor. Thus the logarithm of the partition function was determined only to within an additive constant. That meant that a term proportional to T appeared in the expression for the free energy, or an additive constant appeared in the entropy. This constant, which was sometimes called the chemical constant, could be completely evaluated only after the quantum-mechanical solution was worked out.

10-3 QUANTUM - MECHANICAL EFFECTS

We have mentioned that there are some cases in which the classical approach is not adequate. For these cases it is necessary to include changes in the potential function which result from the motion along the "path." In this section we shall calculate the first-order effect of the potential when the motion of the particle is taken into account.

Instead of approximating $V(x)$ by the constant value $V(x_1)$ in the expression for the density matrix, Eq. (10-43), we might try a Taylor series expansion for V around the point x_1. However, we would find that we could save effort and increase our accuracy if we chose to expand about the mean position given by

$$\bar{x} = \frac{1}{\beta\hbar}\int_0^{\beta\hbar} x(u)\,du \tag{10-50}$$

which is defined for any particular path. We can characterize each path by its mean position and carry out integrations over all such positions instead of integrations over initial positions x_1, as was done in Eq. (10-48). In this way the partition function becomes

$$Z = \int d\bar{x}\int_{x_1}^{x_1}\left[\exp\left(-\frac{1}{\hbar}\left\{\frac{m}{2}\int_0^{\beta\hbar}\dot{x}^2\,du + \int_0^{\beta\hbar} V[x(u)]\,du\right\}\right)\right]\mathfrak{D}x(u) \tag{10-51}$$

In this expression the paths are chosen to satisfy two conditions: (1) that $\bar{x}$ given by Eq. (10-50) is fixed and (2) that the initial and final points are identical (which implies that the integral over all paths also includes an integration over all end points x_1).

Using a Taylor series expansion for $V(x)$ about the point $\bar{x}$, we find

$$\int_0^{\beta\hbar} V[x(u)]\,du = \beta\hbar V(\bar{x}) + \int_0^{\beta\hbar}[x(u) - \bar{x}]V'(\bar{x})\,du + \frac{1}{2}\int_0^{\beta\hbar}[x(u) - \bar{x}]^2 V''(\bar{x})\,du \tag{10-52}$$

By virtue of Eq. (10-50) the second term on the right-hand side of this last equation is zero. Thus, by expanding about the mean position, we arrive at an expression for which the first nonzero correction term is of second order. Using this expansion and including no terms of higher order than the second, we have for the partition function

$$Z \simeq \int e^{-\beta V(\bar{x})}\,d\bar{x}\int_{x_1}^{x_1}\left[\exp\left(-\int_0^{\beta\hbar}\left\{\frac{m}{2}\dot{x}^2 + [x(u) - \bar{x}]^2 V''(\bar{x})\right\}\frac{du}{\hbar}\right)\right]\mathfrak{D}x(u) \tag{10-53}$$

The path integral in this expression differs from those of our previous experience in one particular way. The paths over which the integral is to be evaluated are constrained by Eq. (10-50), which can be rewritten for present purposes as simply

$$\int_0^{\beta\hbar}(x - \bar{x})\,du = 0$$

The substitution $y = x - \bar{x}$ as the path coordinate then gives the constraint in the form

$$\int_0^{\beta\hbar} y\,du = 0$$

and the path integral itself is

$$\int_{x_1-\bar{x}}^{x_1-\bar{x}} \left(\exp\left\{-\int_0^{\beta\hbar}\left[\frac{m}{2}\dot{y}^2 + \frac{1}{2}y^2V''(0)\right]\frac{du}{\hbar}\right\}\right)\mathfrak{D}y(u) \tag{10-54}$$

The integrand of this path integral is the same as that for a harmonic oscillator, with the frequency given by $\omega^2 = -V''(0)/m$.

We now apply the constraint to this path integral in the following way. We multiply the whole path integral by the delta function $\delta\left(\int_0^{\beta\hbar} y\,du\right)$. In order to manipulate the δ function within the path integral, we express it by its Fourier transform,

$$\delta(x) = \int_{-\infty}^{\infty} [\exp(ikx)]\frac{dk}{2\pi}$$

and write

$$\int_{-\infty}^{\infty}\frac{dk}{2\pi}\int_{x_1-\bar{x}}^{x_i-\bar{x}}\left\{\exp\left[-\frac{1}{\hbar}\int_0^{\beta\hbar}\left(\frac{m}{2}\dot{y}^2 + \frac{1}{2}V''y^2 + iky\right)du\right]\right\}\mathfrak{D}y(u) \tag{10-55}$$

In this form, the path integral contains the constraint of Eq. (10-50), and we can therefore proceed directly with standard path integral techniques to obtain the desired solution. We note that the integrand of the path integral now has the same form as the path integral for a forced harmonic oscillator if we interpret both m and V'' to be imaginary. However, we are interested only in the case V'' small, and the approximation of including only the first-order term V'' can be made at any convenient stage.

Problem 10-2 Use the methods of Chap. 3 and, in particular, Eq. (3-66) to solve this path integral. Remember that paths of interest in this problem have the same initial and final points and that completion of the path integral requires an integration over all values of this point. Finally, carry out the integration over all values of k to get as a solution

$$\text{Const}\left[1 - \frac{\beta^2\hbar^2}{24m}V''(\bar{x})\right] \qquad \text{to first order in } V'' \tag{10-56}$$

The partition function which results from the solution obtained in

Prob. 10-2 is written best in the form (valid to first order in V'')

$$Z = \sqrt{\frac{mkT}{2\pi\hbar^2}} \int \left(\exp \left\{ -\beta \left[V(\bar{x}) + \frac{\beta\hbar^2}{24m} V''(\bar{x}) \right] \right\} \right) d\bar{x} \tag{10-57}$$

Here the unknown constant has been evaluated simply by comparison with the classical result of Eq. (10-48). We see from this result that the partition function has the same form which we derive under classical assumptions. The only difference is that a corrective term has been added to the potential. This corrective term, $(\beta\hbar^2/24m)V''(\bar{x})$, is clearly quantum-mechanical in nature, as can be seen from the inclusion of Planck's constant $\hbar$.

Problem 10-3 Show that for many particles (which we identify by subscripts so that the mass of the ith particle is m_i) moving in three dimensions, the correction to the potential is

$$\frac{\beta\hbar^2}{24} \sum_i \frac{1}{m_i} \nabla_i^2 V \tag{10-58}$$

In practice, the results of this calculation are not very useful. In most problems, e.g., in a gas of colliding molecules, the potential rises very sharply so that there is a violent repulsion at small distances. In such a case the second derivative is very large. When this is not so, the formula may be of some use. It has one advantage in that it may be easily extended to another order of accuracy.

Problem 10-4 Show that the correction to the partition function up to the order of $\hbar^4$ contains the factor

$$\left\{1 - \frac{\beta^2\hbar^2}{24m} V''(\bar{x}) + \frac{7\beta^4\hbar^4}{8 \times 720m^2} [V''(\bar{x})]^2 \right.$$
$$\left. - \frac{\beta^3\hbar^3}{24 \times 48m^2} V''''(\bar{x}) + \cdots \right\}$$

We have seen above that quantum-mechanical effects might be represented by calculating the partition function as in the classical formula of Eq. (10-48), but instead of using the correct potential $V(x)$, we use a modified potential $V + \frac{\beta\hbar^2}{24m} V''$. This suggests that we try to go further and seek some possibly better effective potential $U(\bar{x})$ which, when substituted for the true V in the classical equation (10-48), would represent a good approximation to the correct quantum-mechanical partition function. We start out with the exact expression

$$Z = \int \{\exp[-\beta V(\bar{x})]\}\, d\bar{x} \int \exp\left(\frac{m}{2\hbar}\int_0^{\beta\hbar} \dot{x}^2\, du - \frac{1}{\hbar}\int_0^{\beta\hbar} \{V[x(u)] - V[\bar{x}]\}\, du\right) \mathfrak{D}x(u) \qquad \text{(10-59)}$$

We can look upon the path integral in this expression as representing the average over the paths $x(u)$ of the function e^f with

$$f = -\int_0^{\beta\hbar} \{V[x(u)] - V[\bar{x}]\} \frac{du}{\hbar} \qquad \text{(10-60)}$$

where the averaging process is carried out with the weighting function $\exp[-(m/2\hbar)\int \dot{x}^2\, du]$. If we were to replace this average of an exponential with the exponential of an average, thus

$$\langle e^f \rangle \rightarrow e^{\langle f \rangle} \qquad \text{(10-61)}$$

we know we would make an error of the second order in f or, better, of the order of the difference between $\langle f \rangle^2$ and $\langle f^2 \rangle$. We shall see in Chap. 11 that we can determine the sign of this error, i.e., the left-hand side is greater than the right.

We shall obtain an average value of f for each value of $\bar{x}$, that is,

$$\langle f \rangle = \frac{1}{\hbar} \int \left[\exp\left(-\frac{m}{2\hbar}\int_0^{\beta\hbar} \dot{x}^2\, du\right)\right] \int_0^{\beta\hbar} \{V[x(t)] - V[\bar{x}]\}\, dt\, \mathfrak{D}x(u) \qquad \text{(10-62)}$$

where we constrain the path to return to its starting point, and to satisfy the integral restraint given by the definition of $\bar{x}$ in Eq. (10-50). To evaluate this path integral, we define a somewhat different, but related, path integral

$$I(\bar{x}) = \iint \left[\exp\left(-\frac{m}{2\hbar}\int_0^{\beta\hbar} \dot{y}^2\, du\right)\right] \{V[\bar{x} + y(t)] - V[\bar{x}]\}\, \mathfrak{D}y(u)\, dY \qquad \text{(10-63)}$$

where the paths $y(u)$ are subject to the restriction

$$y(0) = y(\beta\hbar) = Y \qquad \int_0^{\beta\hbar} y(u)\, du = 0 \qquad \text{(10-64)}$$

Now, it happens that $I(\bar{x})$ is independent of t. We can see this with the help of the following argument. Suppose each path in the integral is not of finite length, but is really a $\beta\hbar$-length segment of a

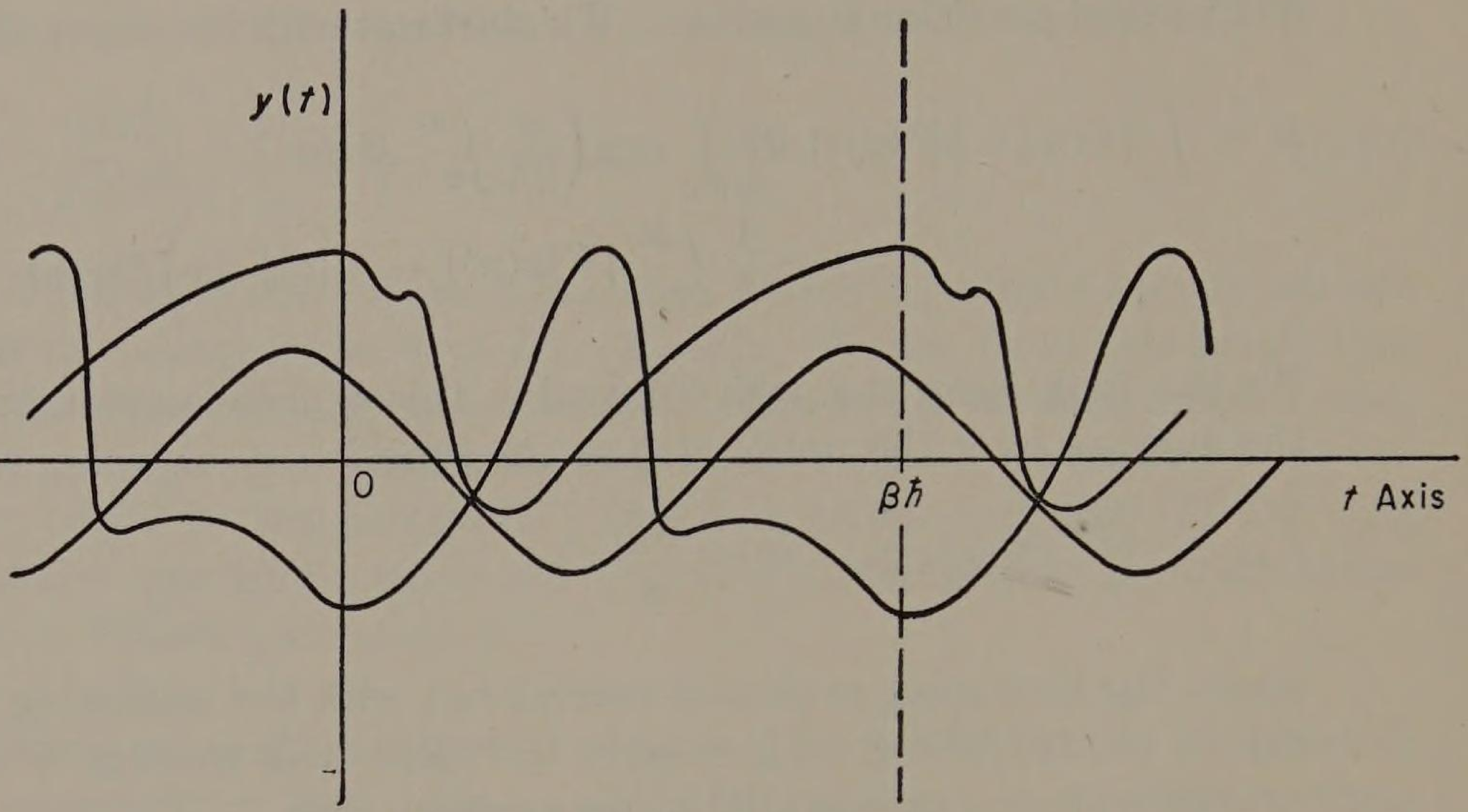


Fig. 10-1 All paths which return at $t = \beta\hbar$ to their initial value (at $t = 0$) can be considered as $\beta\hbar$-length segments of periodic paths where the period is $\beta\hbar$.

periodic path whose period is also $\beta\hbar$, as shown in Fig. 10-1. Consider two of the families of all such paths, one $y(t)$ and the other $y(t_1 + t) = y_1(t)$, as shown in Fig. 10-2. The value which the first attains at $t = t_1$, namely, $y(t_1)$, is reached by the second when its argument is 0, that is, $y_1(0) = y(t_1)$. Furthermore, for any other point

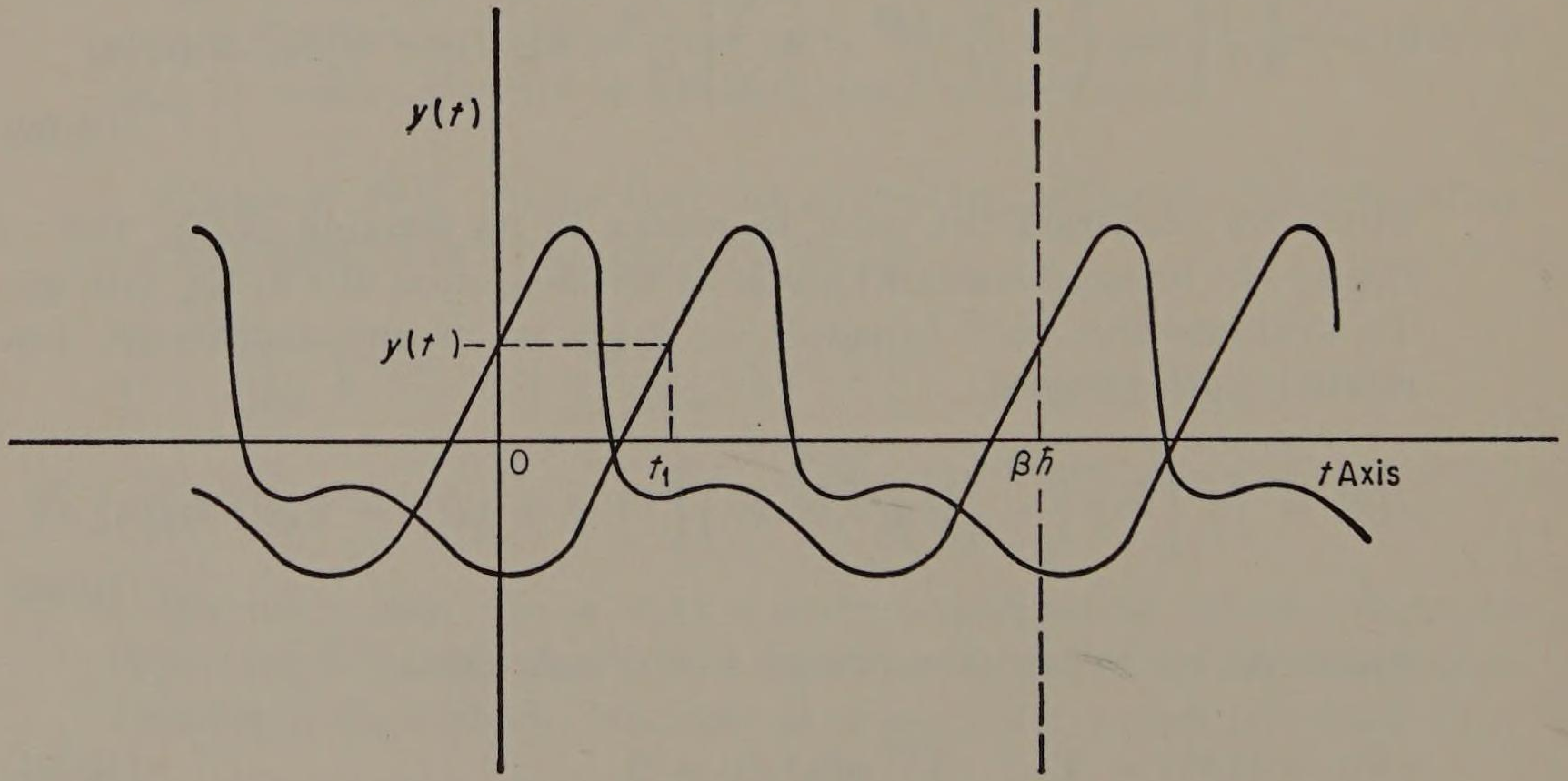


Fig. 10-2 Suppose one of the "periodic" paths $y(t)$, as shown in Fig. 10-1, has the value $y(t_1)$ at $t = t_1$. Then the collection of all "periodic" paths must contain this same path slipped left a distance t_1, that is, $y(t + t_1)$, which will have this same value at $t = 0$. The result of a path integral average over all such paths must then be independent of the selection of the initial point on the t axis.

t_i there exists in the family the analogous function $y_i(t)$ for which $y_i(0) = y(t_i)$, and all such paths give the same contribution to

$$\int_0^{\beta\hbar} y^2\,du$$

Of course, all these statements apply to each path included in the path integral. Thus we see that we lose nothing by arbitrarily setting $t = 0$ in the path integral over all paths $y(u)$, which is the same as saying that the integral is independent of t.

Problem 10-5 Using the method outlined for Prob. 10-2, show that $\langle f \rangle$ and $I(\bar{x})$ are related by

$$I(\bar{x}) = \sqrt{\frac{12m}{2\pi\beta\hbar^2}} \int_{-\infty}^{\infty} [V(\bar{x} + Y) - V(\bar{x})] e^{-6Y^2m/\beta\hbar^2}\,dY = \frac{\langle f \rangle}{\beta} \qquad (10\text{-}65)$$

Suppose we call our approximation to the partition function Z' and call the Helmholtz free energy to which it belongs F', so $Z' = e^{-\beta F'}$. Using the results of Prob. 10-5 and Eq. (10-61) we have

$$Z' = \int (\exp\{-\beta[V(\bar{x}) + I(\bar{x})]\})\,d\bar{x} \int \left[\exp\left(-\frac{m}{2\hbar}\int_0^{\beta\hbar} \dot{x}^2\,du\right)\right] \mathfrak{D}x(u) \qquad (10\text{-}66)$$

The path integral appearing in Eq. (10-66) is the integral evaluated in Eq. (10-46). Thus, we can write

$$e^{-\beta F'} = \sqrt{\frac{mkT}{2\pi\hbar^2}} \int e^{-\beta U(\bar{x})}\,d\bar{x} \qquad (10\text{-}67)$$

where

$$U(\bar{x}) = \int_{-\infty}^{\infty} V(\bar{x} + y) e^{-6y^2m/\beta\hbar^2}\,dy \sqrt{\frac{12m}{2\pi\beta\hbar^2}} \qquad (10\text{-}68)$$

and the term $V(\bar{x})$ has canceled out.

These results mean that we can calculate an approximate free energy F' in a classical manner [i.e., with an expression like Eq. (10-48)] and get a good approximate result if we use an effective potential $U(\bar{x})$, as defined by Eq. (10-68), in place of $V(\bar{x})$. Incidentally, we note that the effective potential is temperature-dependent.

The effective potential is a mean value of $V(x)$ averaged over points near $\bar{x}$ in a gaussian fashion where the mean-square spread (or standard deviation) of the gaussian weighting function is $(\beta\hbar^2/12m)^{1/2}$. Furthermore, if we follow through the various inequalities which are involved in our approximation, we find that the approximate free energy F' exceeds the true free energy F. The details of this are discussed in the next chapter, Eq. (11-9) and following.

Problem 10-6 Show that the relation of Eq. (10-68) becomes the "corrected" potential of Eq. (10-57) (that is, the argument of the exponent in that equation) if V is expanded as a Taylor series.

Problem 10-7 Test the validity of the approximation as it applies to the harmonic oscillator, for which the exact value of the free energy is

$$F_{\text{exact}} = kT \ln \left(2 \sinh \frac{\hbar\omega}{2kT}\right) \tag{10-69}$$

Evaluate the approximate value for the free energy by means of the approximate potential U. Show that

$$U = \frac{m\omega^2}{2}\left(x^2 + \frac{\beta\hbar}{12m}\right) \tag{10-70}$$

and that

$$F_{\text{approx}} = kT\left(\ln \frac{\hbar\omega}{kT}\right) + \frac{(\hbar\omega)^2}{24kT} \tag{10-71}$$

Determine the free energy or, better, the ratio of the free energy to kT, for various values of frequency. It is suggested that the values of 1, 2, and 4, for example, be used for the ratio $\hbar\omega/kT$. Show that F' is greater than F as expected and that the errors grow as the temperature falls. Note that if we are even very far from the classical region (e.g., when the ratio $\hbar\omega/kT = 2$, so that the system has an 85 per cent probability of being in the ground state) the approximate results are still surprisingly close to the true results.

Compare these results with those obtained with the classical approximation in which the free energy is given by $kT \ln (\hbar\omega/kT)$. These results should show the values of $2F/\hbar\omega$ as given in the accompanying table.

$\hbar\omega/kT$	1	2	4
Exact.	0.08266	0.8546	0.9906
Present approx.	0.08333	0.8598	1.0264
Classical.	0.00000	0.6931	0.6931

10-4 SYSTEMS OF SEVERAL VARIABLES

If a system has several variables, the formulas describing them are obtained by direct extension of the methods we have already studied, except for some special problems which arise from consideration of symmetry properties.

Liquid Helium. As an example consider the problem of finding the partition function for liquid helium. Suppose we have N identical atoms, each of mass m, confined in some volume. Suppose further that atoms interact in pairs through a potential $V(r_{1,2})$. This potential is a weak attraction at large distances and a very strong repulsion at short distances. Just to orient our thinking, we might imagine $V(r)$ as the potential for hard spheres. That is,

$$\begin{aligned} &V(r) = 0 \qquad r > a \\ &V(r) = \infty \qquad r < a \qquad a = 2.7\ \text{Å} \end{aligned} \tag{10-72}$$

The lagrangian for such a system has the form

$$L = \sum_i \frac{1}{2} m|\dot{\mathbf{R}}_i|^2 - \frac{1}{2}\sum_{i,j} V(r_{i,j}) \tag{10-73}$$

which means that the partition function is

$$\begin{aligned} Z = \int d^N\mathbf{R}(0) \int \Bigg[\exp\Bigg(-\Bigg\{\frac{m}{2\hbar}\int_0^{\beta\hbar} |\dot{\mathbf{R}}(t)|^2\,dt \\ + \frac{1}{2\hbar}\sum_{i,j}\int_0^{\beta\hbar} V[\mathbf{R}_i(t) - \mathbf{R}_j(t)]\,dt\Bigg\}\Bigg)\Bigg]\,\mathfrak{D}^N\mathbf{R}(t) \end{aligned} \tag{10-74}$$

In this last expression the symbol $d^N\mathbf{R}$ stands for $d^3\mathbf{R}_1\,d^3\mathbf{R}_2\,d^3\mathbf{R}_3 \cdots d^3\mathbf{R}_N$. Similarly, $\mathfrak{D}^N\mathbf{R}$ means $\mathfrak{D}\mathbf{R}_1\,\mathfrak{D}\mathbf{R}_2\,\mathfrak{D}\mathbf{R}_3 \cdots \mathfrak{D}\mathbf{R}_N$. We assume that all the path integrals are taken between initial points $\mathbf{R}_i(0)$ and final points $\mathbf{R}_i(\beta)$ such that $\mathbf{R}_i(0) = \mathbf{R}_i(\beta)$.

The form which we have written down in Eq. (10-74) is actually not correct. The symmetry properties which we mentioned above will affect this result. This characteristic is one of the interesting features of quantum mechanics of identical particles. In Chap. 1 we mentioned that if an event occurs in two indistinguishable ways, then the amplitudes for the two ways will add. In particular, when we are dealing with indistinguishable particles, one alternative way for accomplishing any event always exists; namely, the interchange of two particles. In such a case the amplitudes for the particles (1) as interchanged and (2) as not interchanged must be added. (This addition rule applies to Bose particles. For Fermi particles the contributions for amplitudes which arise from odd permutations of particles will subtract from each other.) Ordinary helium atoms are of isotopic mass 4 and contain 6 particles: 2 protons, 2 neutrons, and 2 electrons. This means that helium atoms are Bose particles and the amplitudes for interchange of particles add. (For instance, we say that Bose particles follow symmetrical statistics, whereas Fermi particles follow antisymmetrical statistics.)

To see how this addition of amplitudes comes about, at least for helium atoms, we can follow this line of argument: In the final state the atoms cannot be distinguished from each other. Thus, although the appearance of the configuration of atoms may be the same finally as it was initially, the identity of some of the atoms may have been exchanged.

For example, an atom which we shall designate as 1 starts a position $x_1(0)$. We have assumed that some atom at least will be in this same position at the close. Thus, for some atom $x(\beta)$ is equal to $x_1(0)$. However, it may not be atom 1 which ends up in this particular place. Instead, atom 1 may go to the initial position of atom 2, say $x_2(0)$, while at the same time atom 2 has moved into the initial position of atom 1. That is, it is possible that atoms 1 and 2 exchange places in the final configuration.

To describe this situation in the most general terms, let Px_i stand for some permutation among the atoms which are initially at x_i. Thus, for example, in the situation in which atoms 1 and 2 were exchanged and all others remained where they were, we would have

$$Px_1 = x_2 \qquad Px_2 = x_1 \qquad Px_3 = x_3 \qquad \cdots \qquad Px_N = x_N \qquad \cdots \tag{10-75}$$

In general, the final state can be any permutation of the initial state:

$$x_j(\beta) = Px_i(0) \tag{10-76}$$

Thus in order to construct the complete amplitude, we must sum amplitudes over all the $N!$ possible permutations, since each permutation represents an alternative possibility. The normalization is correct if we average over all the permutations. The resulting rules for symmetrical statistics mean that Eq. (10-74) must be replaced by

$$Z = \frac{1}{N!} \sum_P{}' d^N\mathbf{R}(0) \int_{\mathbf{R}_i(0)}^{P\mathbf{R}_i(0)} \left[\exp\left(-\frac{1}{2\hbar} \left\{ m \sum_i \int_0^{\beta\hbar} |\dot{\mathbf{R}}_i|^2 \, dt \right.\right.\right.$$
$$\left.\left.\left. + \sum_{i,j} \int_0^{\beta\hbar} V[\mathbf{R}_i(t) - \mathbf{R}_j(t)] \, dt \right\} \right) \right] \mathfrak{D}^N\mathbf{R}(t) \quad (10\text{-}77)$$

where $\sum_P{}'$ means a sum over all permutations P.

If we were dealing with Fermi particles, e.g., the isotope of helium which has three nucleons, we would have to include an extra factor of ± 1, positive for even permutations and negative for odd permutations. There would also be some extra features which depend upon the spin of the atom in our result.

It is possible to give a more detailed derivation of Eq. (10-77) in the following manner. For helium-4 atoms the quantum-mechanical amplitude for two atoms which start at positions a and b to get to positions c and d is

$$K(c,a;d,b) + K(d,a;c,b) \quad (10\text{-}78)$$

(Amplitudes for alternative final conditions add, since these conditions cannot be distinguished from each other.) In this expression $K(c,a;d,b)$ is the complex amplitude to go from a to c while the other particle is going from b to d.

Since the particles are indistinguishable, their symmetry properties imply that the amplitude to find the two particles eventually at the points c and d must be a symmetric function of c and d. That is, the wave function $\psi(c,d)$ must be a symmetric function of the variables r_c, r_d. That is,

$$\psi(c,d) = \psi(d,c) \quad (10\text{-}79)$$

If the particles were Fermi, the wave function would have to be an antisymmetric function of these positions.

If many particles are involved, the rule is simply extended, that is,

$$\begin{aligned} \psi(1,2,3, \ldots ,N) &= \psi(1,3,2, \ldots ,N) \\ &= \psi(1,2,4, \ldots ,N) \\ &= \text{etc.} \end{aligned} \quad (10\text{-}80)$$

The simplest statement of the general rule is that the wave function must be symmetric (antisymmetric for Fermi particles). Although other solutions of Schrödinger's wave equation exist, only symmetric and antisymmetric ones appear in nature. Hence in the sum defining the partition function in Eq. (10-2), we do not wish the sum over *all* energy eigenvalues of the hamiltonian H which can be obtained from solution of $H\phi_n = E_n\phi_n$, but only over those for which the wave function ϕ_n is a symmetric function. For example, the density matrix $\rho(x',x)$ is defined by Eq. (10-28) with a disregard for the statistics of the N atoms involved. How can we reduce this sum to include only symmetric wave functions?

To accomplish this reduction, we use the following trick. First we notice that from any function a symmetric function can be produced simply by permuting all variables and adding together the resulting functions. Thus, for any function $f(x_1,x_2)$ the combination $f(x_1,x_2) + f(x_2,x_1)$ is a symmetric function. It follows that for any wave function $\phi(x_1,x_2, \ldots ,x_N)$ the function

$$\phi'(x_i) = \sum_P \phi(Px_i) \tag{10-81}$$

is symmetrical. Now if $\phi_n(x_i)$ is a solution of the Schrödinger equation, then $\phi'_n(x_i)$ as defined by Eq. (10-81) is also a solution, since the hamiltonian H is symmetric for an interchange of coordinates. Therefore, each interchanged form $\phi_n(Px)$ is a solution, as is the sum.

Some of the energy eigenvalues E_n have eigenfunctions ϕ_n which are symmetric, and some do not. Suppose E_k is an energy eigenvalue for which the Schrödinger equation does not have a symmetric solution. Then the sum $\sum_P \phi_k(Px)$ must vanish, since if it existed it would be a symmetric solution for E_k. This result implies that the operation defined by Eq. (10-81) selects just those solutions to the wave equations which are symmetric. All other solutions vanish. If $\phi_n(x)$ is symmetric, then it is equal to $\phi_n(Px)$; and since there are $N!$ ways of permuting the N atom, we have

$$\begin{aligned} \sum_P \phi_n(Px_i) &= N!\phi_n(x_i) && \text{if } \phi_n \text{ is symmetric} \\ &= 0 && \text{if } \phi_n \text{ is of any other symmetry} \end{aligned} \tag{10-82}$$

These results give us an answer to our question. We can now select out of the sum defining the density matrix those particular elements which apply to symmetric states. Thus

$$\begin{aligned} \sum_P \rho(Px',x) &= \sum_n^{\text{all}} \sum_P \phi_n(Px')\phi_n^*(x)e^{-\beta E_n} \\ &= N! \sum_n^{sym} \phi_n(x')\phi_n(x)e^{-\beta E_n} \\ &= N!\rho_{sym}(x',x) \end{aligned} \tag{10-83}$$

This is the reason why in Eq. (10-77) defining the partition function for symmetric statistics we permute all the particles and divide by $N!$. The resulting partition function corresponds to

$$\int \rho_{sym}(x_0,x_0)d^Nx_0 = Z_{sym} = \sum_n^{sym} e^{-\beta E_n} \tag{10-84}$$

We note some of the features of Eq. (10-77). At high temperatures, we should expect a classical solution for the partition function with no quantum-mechanical effects in evidence. Suppose we disregard the effects of the potential for the moment and consider the effect of the motion of an atom from its initial point to some other point a distance d away. In the path integral of Eq. (10-77), this is a motion from the initial point $\mathrm{R}_i(0)$ to the permuted position $P\mathrm{R}_i(0)$, and the contribution of that particular permutation to the sum over all permutations is proportional to $\exp(-md^2kT/2\hbar^2)$, thus decreasing with increasing temperature or increasing spacing between the atoms. Hence, unless the atoms are extremely close together, no permutation in the sum is important—even the simplest interchange between two atoms—in comparison with the identity permutation which leaves all atoms in their original location. If we include the effects of the potential which increases steeply at a radius of 2.7 Å from the center of an atom in liquid helium, then no configurations in which the atomic spacing is less than this value are important.

Since only the identity permutation makes a significant contribution to the summations, all that remains for our consideration is the factor $1/N!$. In the early days of *classical* thermodynamics it was realized that such a factor was convenient when dealing with identical particles, but its significance was not clearly understood. Its effect on the chemical constant is called the *entropy of mixing* when systems of several different kinds of atoms are studied.

As the temperature falls, the exponential factor $\exp(-md^2kT/2\hbar^2)$ prejudicing against migrations to new final positions becomes smaller and smaller. This means that at extremely low temperatures new terms become important in the summation over permutations. Of course, the quantum modifications must be included; and we saw they could be included as a first approximation by replacing the potential V by a corrected potential U. As the temperature falls, the specific heat of liquid helium begins to rise slightly near about 2.3 or 2.4°K.

Problem 10-8 The density of liquid helium is 0.17 g/cm³. Give an order-of-magnitude estimate of the temperature at which perturbation terms should begin to play an important role in the description of liquid helium.

At first sight, one would not expect very elaborate exchanges of atoms to ever be important. An exponential factor involving the spacing must be included each time an atom moves to its neighboring location. If we call this factor y, then for r atoms to move to neighboring spots the factor y^r must be included, and since y is certainly less than 1 at any temperature, y^r could become quite small for large r. We certainly would think that as r approaches any reasonable fraction of the approximately 10^{22} atoms in a cubic centimeter of liquid helium, contributions from factors like y^r must be infinitesimal. However, this first sight does not take into account the fact that with r atoms permuting, there are an enormous number of possible permutations, $r!$. Thus the small weight of one particular permutation is offset by the large number involved.

Another question which arises in the description of liquid helium concerns the type of permutations which are involved. Any permutation can be described by cycles; thus 1–4, 4–7, 7–6, 6–1 is a cycle. Are the important cycles long or short? A careful estimate shows that at moderate temperatures, only simple exchanges of two atoms are important. Then as the temperature falls, cycles of three atoms become important, then four, and so on. But then suddenly, at a certain critical temperature, cycles of much greater length L offset by their great number the small value of y^L. At this temperature cycles of importance become very long, involving nearly all of the atoms inside a container. At this point the curve of specific heat vs. temperature shows a discontinuity. Below this temperature the behavior of the liquid is very strange. It flows through very thin tubes without any resistance for low velocities. It stimulates infinite heat conductivity in bulk, etc. These odd characteristics are manifestations of quantum mechanics, particularly the constructive interference between amplitudes for replacing one atom with another. Quantitatively, the details of the behavior of the specific heat just at the transition temperature are not on a very firm foundation. But the qualitative reason for the transition is clear.[1]

The expression analogous to Eq. (10-77) for Fermi particles, such as He^3, is also easily written down. However, in the case of liquid He^3, the effect of the potential is very hard to evaluate quantitatively in an accurate manner. The reason for this is that the contribution of a cycle to the sum over permutations is either positive or negative depending on whether the cycle has an odd or even number of atoms in its length L. At low temperature, the contributions of cycles such as $L = 51$ and $L = 52$ are very nearly equal but opposite in sign, and therefore they very nearly cancel. It is necessary to compute the difference between such terms, and this requires very careful calculation of each term separately. It is very difficult to sum an alternating series of large terms which are decreasing slowly in magnitude when a precise analytic formula for each term is not available.

---

[1] A more detailed discussion of the partition function of liquid helium from this point of view may be found in R. P. Feynman, *Phys. Rev.*, vol. 91, p. 1291, 1953.

Progress could be made in this problem if it were possible to arrange the mathematics describing a Fermi system in a way that corresponds to a sum of positive terms. Some such schemes have been tried, but the resulting terms appear to be much too hard to evaluate even qualitatively.

For molecules which are separated by distances in the neighborhood of 1 Å we have seen that the effects of exchange (the nonidentical permutations) are important only when the temperature is down to a few degrees absolute. In contrast to this, consider the behavior of electrons in a solid metal. The mass of the electron is so much smaller than that of a molecule that the critical temperature is much higher. At room temperatures, electrons in a metal are described accurately only by equations which include the exchange effects of these cyclic permutations. From this point of view, room temperature is very cold for electrons. The exchange effects are of dominant importance, or, to put it another way, the electron gas is degenerate. Of course, the electrons interact by Coulomb's law, which is quite strong. But since the effects of the Coulomb attraction are of long range, they tend to average out. To a fair approximation, the electrons act as if they were independent, although, of course, each moves in the same periodically varying potential produced by the arrangement of the nuclei and the average of the positions of neighboring electrons. From the study of the ideal Fermi gas neglecting interactions, we can learn a lot about the behavior of electrons in metals.

However, it is apparent that we cannot learn quite enough, for the superconductivity of metals occurring below a few degrees absolute would remain a mystery. This phenomenon, in some metals at least, involves an interaction in which the slow vibratory motion of the atoms is involved. We conclude this because the transition temperatures for two different isotopes of the same metal depend on the atomic mass. This value of the isotopic mass would not be important if the transition were simply a matter of mutual interaction between electrons, or interaction of the electrons with an idealized array of *fixed* atoms. The idealization that the atoms are fixed must be incorrect. But how does the motion of the atoms produce a sudden jump in specific heat in some metals and permit electrical conductivity below this temperature without resistance? This question was first answered in a convincing way by Bardeen, Cooper, and Schrieffer.[1] The path integral approach played no part in their analysis, and in fact it has never proved useful for degenerate Fermi systems.

The Planck Blackbody Radiation Law. The partition function for any system of interacting oscillators is easily worked out. Such a system is equivalent to a set of independent oscillators of frequencies ω_i. However, the value of free energy F for independent systems is the sum of the values of F for each of the separate systems, which we find directly from the sum of Eq. (10-2) to be

$$kT \ln \left(2 \sinh \frac{\hbar\omega}{2kT}\right)$$

This gives the free energy of a linear system as

$$\begin{aligned} F &= kT \sum_i \ln \left(2 \sinh \frac{\hbar\omega_i}{kT}\right) \\ &= kT \sum_i \ln \left(1 - e^{-\hbar\omega_i/kT}\right) + \sum_i \frac{\hbar\omega_i}{2} \qquad (10\text{-}85) \end{aligned}$$

The last term in this expression is a ground-state energy of the system.

For an electromagnetic field in a box of volume V, the modes are defined by the vector wave number $\mathbf{K}$, two for each $\mathbf{K}$. The zero-point energy is omitted. Thus the free energy of the electromagnetic field computed per unit volume is

$$F = kT \int \frac{d^3\mathbf{K}}{(2\pi)^3}\, 2 \ln \left(1 - e^{-\hbar Kc/kT}\right) \qquad (10\text{-}86)$$

The internal energy U is the partial derivative of βF with respect to β which becomes (putting $\omega = Kc$)

$$U = 2 \int \frac{d^3\mathbf{K}}{(2\pi)^3}\, \hbar\omega\, \frac{1}{e^{\hbar\omega/kT} - 1} \qquad (10\text{-}87)$$

The volume element in $\mathbf{K}$ space can be written as

$$d^3\mathbf{K} = 4\pi K^2\, dK = 4\pi \frac{\omega^2}{c^3}\, d\omega \qquad (10\text{-}88)$$

---

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.*, vol. 106, p. 162, 1957, and vol. 108, p. 1175, 1957.

This means that the energy in the electromagnetic field in the range of frequencies from ω to $\omega + d\omega$ is

$$\frac{2 \cdot 4\pi}{(2\pi c)^3} \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1} \tag{10-89}$$

This is the famous blackbody-radiation law discovered by Planck. It was the first real quantitative quantum-mechanical result discovered and was the first step in the discovery of the new laws.

Another early quantum-mechanical triumph was the explanation of the temperature dependence of the specific heat of solids by Einstein and by Debye. This also comes from Eq. (10-85), but the oscillators are now the normal modes of the crystal, as described in Chap. 8. For example, the thermal energy per unit volume of such a crystal is, like Eq. (10-87) (leaving out the zero-point energy), just

$$U = \sum_{3p \text{ modes}} \int \frac{\hbar\omega(\mathbf{k})}{\exp[\hbar\omega(\mathbf{k})/kT] - 1} \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{10-90}$$

where $\omega(\mathbf{k})$ is the frequency of a phonon of wave vector $\mathbf{k}$. In a crystal, this is a multiple-valued function (there are $3p\omega$ values for each $\mathbf{k}$ if there are p atoms in a unit cell), and we must sum over each of the possible values of ω for each $\mathbf{k}$. The $\mathbf{k}$ integral extends only over the finite range proper for the crystal. For light there are two modes for each $\mathbf{k}$, each of frequency $\omega = kc$, so the sum gives a factor 2 and Eq. (10-87) results, the integral on $\mathbf{k}$ now going to infinity.

The result of Eq. (10-90), studied in various approximations by Einstein and Debye, gave a good accounting of the main features of the specific-heat curve, particularly the behavior at low temperatures, which had been in direct contradiction to the classical expectations. Today, putting a more complete knowledge of the phonon spectrum $\omega(\mathbf{k})$ into Eq. (10-90) yields a completely satisfactory description of that part of the specific heat of solids due to internal vibration of the atoms.

10-5 REMARKS ON METHODS OF DERIVATION

The presentation of statistical mechanics given in the early part of this chapter leaves much to be desired. The fundamental law which shows that the probability for finding a system in the energy state E is proportional to $e^{-E/kT}$ is usually derived by considering the interaction of complex systems over long periods of time. But an entertaining problem presents itself.

We have started our discussion of physics in this book with a definition of the laws of quantum mechanics using a path integral technique, as shown in Chap. 2. Just as a question of curiosity, let us take the point of view that this is the fundamental law. Then ultimately these statistical properties of a system whose quantum-mechanical properties are defined by such a path integral are found to be expressible in terms of the partition function Z. This function can be defined by a path integral of an obviously very similar and closely related form, as shown by Eq. (10-77). Yet the derivation of this result requires noting the wave equation, the existence of stationary states and eigenvalues, and the argument about interaction over long periods of time to which we referred, all of which leads to the expression (10-1) for the partition function in terms of the energy levels E_i. Finally, we proceed to the reverse argument producing the path integral formulation for Z. Is there any way to derive the path integral expression for Z for a system in equilibrium directly from the path integral description for the time-dependent motion? Can we find a short cut which avoids the mention of energy levels altogether? If it is possible, we do not yet know how to do it.

One might ask: Why try it at all? It is like showing that you can swim with your hands tied at your back. After all, you know there are energy levels. The only excuse for trying to avoid their mention would be that in so doing a deeper understanding of physical processes might result or possibly more powerful methods of statistical mechanics might be evolved. At any rate, it would be interesting to solve the problem.

It was the promptings of a similar quest, to get the well-known variational principle permitting the computation of the lowest energy level of a system directly from the path integral formulation (instead of indirectly via the Schrödinger equation), which resulted in the methods described in Chap. 11. Thus the results of this apparently academic problem were of some use as well as of some interest.

Nevertheless, if we prefer, we can suppose our desire for one particular course in achieving a solution is prompted simply by an academic interest in the methods of classical physics. Suppose that we have a system obeying the principle of least action, with the action defined by

$$S = \frac{1}{2}\int m\dot{x}^2\,dt + \frac{k}{2}\int x(t)x(t+a)\,dt \tag{10-91}$$

so that the equation of motion is

$$m\ddot{x} = \frac{k}{2}[x(t+a) + x(t-a)] \tag{10-92}$$

Here we have created the curious situation in which a particle is driven by a force depending on the average value of coordinates that were and that will be. There are exponentially exploding solutions of Eq. (10-92), but let us say that only motions for which x remains finite both in the distant past and in the distant future will be allowed. (Incidentally, it is likely that solutions which we wish to ignore are excluded anyway if the action law is stated as $\delta S = 0$ for all variations of path δx subject to the restraint $\delta x \to 0$ for $t \to \pm\infty$.)

For such a system it is possible to define an expression for energy which is conserved; for the equations of motion of the system do not depend on time. (No simple hamiltonian gives the equations of motion.) Presumably, such a system could possess properties which allow it, for example, to be perturbed by molecules of a gas and thus achieve thermal equilibrium. We might ask: What are the averages of various quantities describing a system obeying the equations of motion of Eq. (10-92) and appropriate boundary conditions at infinity when it is in equilibrium at the temperature T? Perhaps such a problem is not definable, or perhaps it is easily solved only in this special case because the equations of motion are linear. But the aim of these remarks is to ask whether the existence of a hamiltonian and momentum variables is indeed necessary to the formulation of classical statistical mechanics—or whether a wider class of mechanical systems can be analyzed, a system in which the equation of motion comes most simply from the principle of least action, even though that action involves more than the instantaneous positions and velocities of the particles in the system.

This question is the classical analogue of our more interesting question, namely, how do we proceed directly from the path integral formulation of quantum-mechanical laws of a mechanical system to the path integral formulation of statistical mechanical laws for the same system in equilibrium?

Problem 10-9 Show that the expression

$$E(t) = \frac{1}{2}m[\dot{x}(t)]^2 - \frac{k}{2}x(t)x(t+a) + \frac{k}{2}\int_t^{t+a} x(t'-a)\dot{x}(t')\,dt' \tag{10-93}$$

defines a conserved energy for the equation of motion (10-92).

In general, for any action functional, like S, that does not involve the time explicitly (i.e., is invariant for the transformation $t \rightarrow t +$ const) there is an expression $E(T)$ for the energy at time T which is conserved. It can be found by asking for the first-order change in the action S when all paths are changed from $x(t)$ to $x[t + \eta(t)]$, where $\eta(t) = +\epsilon/2$ for $t > T$ and $\eta(t) = -\epsilon/2$ for $t < T$, with constant ϵ. δS is then $\epsilon E(T)$ for infinitesimal ϵ.

Problem 10-10 Discuss the problem of the path integral formulation of statistical mechanics for a particle which is in a time-constant magnetic field.

11

The Variational Method

IN this chapter we discuss a method based on a variational principle for the approximate evaluation of certain path integrals. First, we shall illustrate the method by some examples. Later, we consider those problems for which the method may be useful.

11-1 A MINIMUM PRINCIPLE

Suppose we wish to evaluate the free energy F of a system. This problem can be expressed in terms of path integrals by starting with the partition function for the system defined in Eq. (10-4) as

$$Z = e^{-\beta F} \tag{11-1}$$

In Eq. (10-30) the partition function was expressed as an integral over the density matrix $\rho(x,x)$. Then, in Sec. 10-2, a kernel expression for $\rho(x,x)$ was developed. It allows us to write

$$Z = \int_{-\infty}^{\infty} \int_{x_1}^{x_1} e^{S/\hbar}\, \mathfrak{D}x(u)\, dx_1 \tag{11-2}$$

so long as we use the "time" variable u as an imaginary quantity.

In Sec. 10-3 we developed a perturbation technique for the evaluation of the path integral defining the partition function for certain special cases. We shall now describe another technique, applicable in those cases where S *is real*. For ordinary cases without a magnetic field (and no spin) S is real.

Throughout the remainder of this chapter, we assume that we are choosing units in such a way that the value of $\hbar$ is 1. Whenever it is necessary to include $\hbar$ symbolically in order to visualize the quantum-mechanical character of a result, it can be so included by a straightforward dimensional inspection.

Let us suppose that some other S' can be found which satisfies two conditions: First, S' is simple enough that expressions such as $\int e^{S'}\, \mathfrak{D}x(t)$ or $\int G e^{S'}\, \mathfrak{D}x(t)$, for simple functionals G, can be evaluated. Second, the important paths in the integral $\int e^{S}\, \mathfrak{D}x(t)$ and those in the integral $\int e^{S'}\, \mathfrak{D}x(t)$ are similar, that is, S' and S are similar when they are both small. Now suppose F' is the free energy associated with S'. That is,

$$e^{-\beta F'} = \int_{-\infty}^{\infty} \int_{x_1}^{x_1} e^{S'}\, \mathfrak{D}x(u)\, dx_1 \tag{11-3}$$

so that

$$\frac{\int\int e^{S}\, \mathfrak{D}x(t)\, dx_1}{\int\int e^{S'}\, \mathfrak{D}x(t)\, dx_1} = e^{-\beta(F-F')} \tag{11-4}$$

Then since $e^S = e^{S-S'}e^{S'}$, we can write Eq. (11-4) as

$$\int\int e^{S-S'}e^{S'}\,\mathfrak{D}x(t)\,dx_1\left[\int\int e^{S'}\,\mathfrak{D}x(t)\,dx_1\right]^{-1} = e^{-\beta(F-F')} \tag{11-5}$$

This says simply that $e^{-\beta(F-F')}$ is the average value of $e^{S-S'}$ where this average is taken over all paths with the same initial and final point and the weight of each path is $e^{S'}\,\mathfrak{D}x(t)$. All possible values of x_1 are included in the averaging process.

One way to proceed now would be to suppose that $S - S'$ is small and that $F - F'$ is small and then expand both sides up to the first power in their respective exponents. This method appears dubious because $\beta(F - F')$ is not small if β is large. However, comparison of higher-power terms shows that this is nevertheless a legitimate approximation to $F - F'$.

The argument can be made much more rigorous and powerful in the following way. The average value of e^x when x is a random variable always exceeds or equals the exponential of the average value of x, as long as x is real and the weights used in the averaging process are positive. That is,

$$\langle e^x \rangle \geq e^{\langle x \rangle} \tag{11-6}$$

where $\langle x \rangle$ = the weighted average of x. This follows because the curve of e^x is concave upward, as shown in Fig. 11-1, so that if a number of masses (weights) lie along this curve, the center of gravity of these masses lies above the curve. The vertical height of this center of gravity is the average vertical position $\langle e^x \rangle$ of the points. It

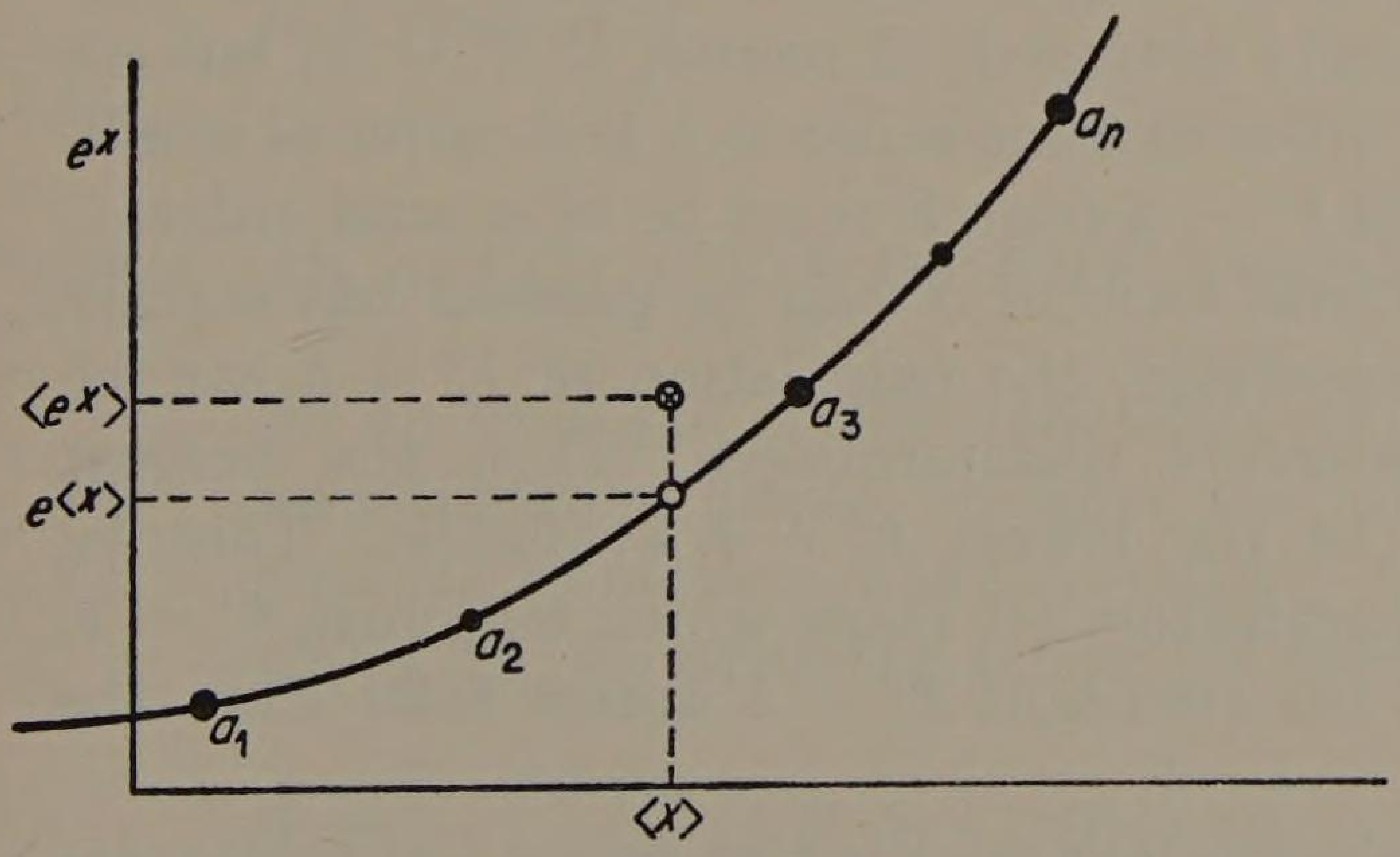


Fig. 11-1 We assume the weighting factors a_i are positive and look on them as different masses positioned along a string. Then the exponential of the weighted average of x, that is, $e^{\langle x \rangle}$, must lie below the weighted average of the exponentials $\langle e^x \rangle$ because of the concave nature of the curve e^x. The value of $e^{\langle x \rangle}$ must lie in the curve, but $\langle e^x \rangle$, the center of gravity of the several points, must lie above the curve.

exceeds $e^{\langle x\rangle}$, the ordinate of the curve e^x at the abscissa position of the center of gravity, which is the average value $\langle x\rangle$.

On the left-hand side of Eq. (11-5) we take the average value of $e^{S-S'}$ over paths with the positive weights $e^{S'}\,\mathfrak{D}x(t)$, where S' and S are real. Hence, by Eq. (11-6), this exceeds the quantity $e^{S-S'}$, where $\langle S-S'\rangle$ is the average of $S-S'$ with this same weighting scheme, namely, with the weight $e^{S'}\,\mathfrak{D}x(t)$. That is,

$$\langle S-S'\rangle = \int\int (S-S')e^{S'}\,\mathfrak{D}x(t)\,dx_1\left[\int\int e^{S'}\,\mathfrak{D}x(t)\,dx_1\right]^{-1} \tag{11-7}$$

We have then

$$e^{\langle S-S'\rangle} \leq e^{-\beta(F-F')} \tag{11-8}$$

This result implies that

$$F_0 \leq F_0' - \frac{1}{\beta}\langle S-S'\rangle \tag{11-9}$$

Our final result is then

$$F \leq F' - \delta \tag{11-10}$$

where

$$\delta = \frac{1}{\beta}\int\int (S-S')e^{S'}\,\mathfrak{D}x(t)\,dx_1\left[\int\int e^{S'}\,\mathfrak{D}x(t)\,dx_1\right]^{-1} \tag{11-11}$$

It is very fortunate that we have a minimum principle here. It says that, if we calculate $F_0' - \delta$ for various "actions" S', that calculation which gives the smallest result is nearest to the true free energy F.† The energy F is actually obtained, of course, if $S' = S$; but we can guess that if S and S' differ in some sense to a first order of smallness, then the derivation of $F' - \delta$ from F must be of second order.

If only a reasonable general form of S' can be guessed but certain parameters still remain uncertain, the calculation of $F' - \delta$ can be made leaving these parameters undetermined. Then the nearest approximation to F will be the lowest $F' - \delta$ available. That is, the "best" values of the parameters are those which minimize $F' - \delta$, "best" in the sense that the resultant $F' - \delta$ differs least from the true F.

This same minimum principle can be used to find an approximate value for the lowest energy state of the system E_0. Recall that

---

† Again, it is worth emphasizing that neither S nor S' is an action functional in the proper physical meaning of the term, since both are defined with the imaginary variable u used as the "time" variable. However, operations with path integrals are the same for these functionals as for proper physical actions used previously.

$$Z = e^{-\beta F} = \sum_{n=0}^{\infty} e^{-\beta E_n} \tag{11-12}$$

As the temperature of the system becomes lower and lower, that is, as β grows larger, terms involving higher values of energy become less and less important in this series. Eventually the series for Z is dominated by the term of smallest energy, $e^{-\beta E_0}$. That is,

$$\lim_{\beta\to\infty} Z = e^{-\beta E_0} \tag{11-13}$$

Now following the line of argument developed in the preceding paragraphs, we can simply replace F with E_0. We define E_0' as the result of the path integral involving the new action S' and finally derive

$$E_0 \leq E_0' - \delta \tag{11-14}$$

as an approximation in the limit of large β.

In approximating E_0 by this technique, our task is somewhat simpler than it was for the free energy F. Specifically, we can disregard the specification that the initial and final points of the paths be the same. To understand this, we refer back to Eq. (10-28) and note that as β becomes large the density matrix $\beta(x',x)$ is also dominated by the zero-order term and approaches $e^{-\beta E_0}\phi_0(x')\phi_0^*(x)$. Thus the dependence on x' and x enters into a multiplying factor but does not affect the nature of the exponential behavior of the function. It is this exponential behavior which is fundamental in the evaluation of E_0 by this technique.

11-2 AN APPLICATION OF THE VARIATIONAL METHOD

As an example of the evaluation of a partition function using this variational principle, consider the example of a single particle constrained to move in one dimension. Using the approach developed in Chap. 10, we write the action for such a particle as

$$S = -\int_0^\beta \left\{\frac{m}{2}\dot{x}(t)^2 + V[x(t)]\right\} dt \tag{11-15}$$

So, for large values of β, the partition function is

$$e^{-\beta E_0} \simeq \iint_{x_0}^{x_0} \left[\exp\left(-\int_0^\beta \left\{\frac{m}{2}\dot{x}^2 + V[x(t)]\right\} dt\right)\right] \mathcal{D}x(t)\, dx_0 \tag{11-16}$$

This path integral is over paths which return to the initial starting points; and after the path integral has been evaluated, a further integration over all possible starting points is carried out.

In Sec. 10-2 we considered this same problem and pointed out how the classical approximation may be derived by inspection. In the classical limit of high temperatures or high values of kT compared with $\hbar$, the value of $\beta\hbar$ is so small that paths which get very far away from x_0 do not contribute. Thus, the potential can be replaced by the constant value $V(x_0)$, and the path integral contributes only a constant, giving

$$e^{-\beta E_0}\,(\text{classical}) = \sqrt{\frac{m}{2\pi\beta}} \int e^{-\beta V(x)}\,dx \tag{11-17}$$

as shown in Eq. (10-48).

In Chap. 10, one quantum-mechanical improvement was made on the classical result by expanding the potential about the average position of the path and using terms up through the second order in this expansion. Then a still greater improvement was achieved by using a potential U, developed in a particular averaging process. From the point of view of this chapter, we see that that approach was a special application of the variational method. To clarify this point, we shall review the key steps using the notation and concepts of this chapter.

Thus we wish to derive a suitable trial function $W(\bar{x})$, where $\bar{x}$ is the average position of a path defined by

$$\bar{x} = \frac{1}{\beta}\int_0^\beta x(t)\,dt \tag{11-18}$$

Along any particular path, this substitute for the potential is a constant, so that the new form of the action along that path becomes

$$S' = -\int_0^\beta \frac{m}{2}\dot{x}^2\,dt - \beta W(\bar{x}) \tag{11-19}$$

With this more general form, it is possible to calculate both F' and $\langle S - S'\rangle$.

Proceeding along this course, we use Eq. (11-14). Substituting into this expression, we have

$$\delta = \frac{\displaystyle -\iint \left\{\frac{1}{\beta}\int_0^\beta V[x(t')]\,dt' - W(\bar{x})\right\} \times \left[\exp\left(-\int_0^\beta \frac{m}{2}\dot{x}^2\,dt\right)\right]\{\exp[-\beta W(\bar{x})]\}\,\mathfrak{D}x(t)\,dx_0}{\displaystyle \iint \left[\exp\left(-\int_0^\beta \frac{m}{2}\dot{x}^2\,dt\right)\right]\{\exp[-\beta W(\bar{x})]\}\,\mathfrak{D}x(t)\,dx_0} \tag{11-20}$$

It is to be remembered that the paths to be used in the path integrals of Eq. (11-20) are those which have the same initial and final points, and, as in Eq. (11-16), a further integration over all end points x_0 is to be carried out.

Note that the numerator of δ is quite similar to the term $I(\bar{x})$ introduced in Eq. (10-63), if we restrict ourselves to paths that have a specific average value $\bar{x}$ and count on integrating over all possible values of $\bar{x}$ at a later stage of the calculation. By the same arguments as were used in the discussion of $I(\bar{x})$, we see that the numerator of δ is independent of t'. We can evaluate the path integrals in both numerator and denominator by the methods used in Chap. 10 and take the answer from Eq. (10-65), remembering that

$$Y = x_0 - \bar{x} \tag{11-21}$$

Since the denominator is simply a special form of the expression appearing in the numerator, the result is

$$\delta = \frac{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [V(x_0) - W(\bar{x})] \left\{\exp\left[-\frac{6m}{\beta}(x_0 - \bar{x})^2\right]\right\} \times \{\exp[-\beta W(\bar{x})]\}\, dx_0\, d\bar{x}}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\{\exp\left[-\frac{6m}{\beta}(x_0 - \bar{x})^2\right]\right\} \times \{\exp[-\beta W(\bar{x})]\}\, dx_0\, d\bar{x}} \tag{11-22}$$

The integral over x_0 in the denominator in Eq. (11-22) can be easily evaluated to give $(\beta\pi/6m)^{1/2}$. Furthermore, the integral over the term in the numerator containing the factor $W(\bar{x})$ results in this same multiplying constant. It will be more convenient for our future work if we carry out that particular integration in the numerator and further simplify the resulting expression by defining the function $\overline{V(\bar{x})}$ as

$$\overline{V(\bar{x})} = \sqrt{\frac{6m}{\pi\beta}} \int_{-\infty}^{\infty} V(x_0) \left\{\exp\left[-\frac{6m}{\beta}(x_0 - \bar{x})^2\right]\right\} dx_0 \tag{11-23}$$

The form of $\overline{V(\bar{x})}$ reveals the quantum-mechanical effect we have introduced. This function is a weighted average of $V(x_0)$ with a gaussian weighting function just like the function $U(x_0)$ defined by Eq. (10-68), and the gaussian spread is again $(\beta\hbar^2/12m)^{1/2}$. For a helium atom at a temperature of 2°K, this spread amounts to about 0.7 Å. At room temperatures, however, it is only about 2 per cent of the 2.7-Å diameter of the atom. The value of δ can now be written as

$$\delta = \frac{\int [W(\bar{x}) - \overline{V(\bar{x})}]\{\exp[-\beta W(\bar{x})]\}\,d\bar{x}}{\int \{\exp[-\beta W(\bar{x})]\}\,d\bar{x}} \tag{11-24}$$

The next step is to evaluate $W(\bar{x})$ by the requirement that we obtain the minimum value for $F' - \delta$, as shown in Eq. (11-13). F' is defined by

$$\begin{aligned}\exp(-\beta E_0') &= \int e^{S'}\mathfrak{D}x(t)\\ &= \int \left\{\exp\left[-\int_0^\beta \frac{m}{2}\dot{x}^2\,dt - \beta W(\bar{x})\right]\right\}\mathfrak{D}x(t)\\ &= \int \{\exp[-\beta W(\bar{x})]\}\\ &\qquad \times \int_{\bar{x}\text{ fixed}} \left[\exp\left(-\int_0^\beta \frac{m}{2}\dot{x}^2\,dt\right)\right]\mathfrak{D}x(t)\,d\bar{x}\end{aligned} \tag{11-25}$$

The path integral is a simple one whose value we know to be $\sqrt{m/2\pi\beta}$, so that we obtain

$$\exp(-\beta E_0') = \sqrt{\frac{m}{2\pi\beta}} \int \{\exp[-\beta W(\bar{x})]\}\,d\bar{x} \tag{11-26}$$

The next step, finding the optimum choice for $W(\bar{x})$, requires us to determine the effect of a small variation in the function $W(\bar{x})$ on the value of $F' - \delta$ and set this effect equal to 0. Thus, imagining W to be replaced by

$$W \to W(\bar{x}) + \eta(\bar{x}) \tag{11-27}$$

we find from Eq. (11-26) that the variation in F' is

$$\partial E_0' = \frac{\int \eta(\bar{x})\{\exp[-\beta W(\bar{x})]\}\,d\bar{x}}{\int \{\exp[-\beta W(\bar{x})]\}\,d\bar{x}} \tag{11-28}$$

and the variation in δ from Eq. (11-24) is

$$\begin{aligned}\partial\delta &= \frac{\int \{\exp[-\beta W(\bar{x})]\}\{\beta\eta(\bar{x})[\overline{V(\bar{x})} - W(\bar{x})] + \eta(\bar{x})\}\,d\bar{x}}{\int \{\exp[-\beta W(\bar{x})]\}\,d\bar{x}}\\ &\quad + \frac{\int \{\exp[-\beta W(\bar{x})]\}[W(\bar{x}) - \overline{V(\bar{x})}]\,d\bar{x} \times \int \beta\eta(\bar{x})\{\exp[-\beta W(\bar{x})]\}\,d\bar{x}}{(\int \{\exp[-\beta W(\bar{x})]\}\,d\bar{x})^2}\end{aligned} \tag{11-29}$$

Finding a stationary value for the right-hand side of Eq. (11-13) requires simply that

$$\partial E_0' - \partial\delta = 0 \qquad (11\text{-}30)$$

which will be true if we take

$$W(\bar{x}) = \overline{V(\bar{x})} \qquad (11\text{-}31)$$

This, in turn, implies that δ is 0 and that F' has the same form as the classical free energy defined in Eq. (11-17). However, the potential, in the definition of F', has been replaced by $V(\bar{x})$. That is,

$$\exp(-\beta E_0') = \sqrt{\frac{m}{2\pi\beta}} \int \{\exp[-\beta\overline{V(\bar{x})}]\}\, d\bar{x} \qquad (11\text{-}32)$$

where $\overline{V(\bar{x})}$, the effective classical potential, is given by Eq. (11-24). For large values of β, the free energy is essentially the same as the lowest energy level E_0; thus we can interpret Eq. (11-32) as providing an approximation to E_0. This means that the variational approach has produced the same result as that obtained in Chap. 10 and shown in Eqs. (10-67) and (10-68).

11-3 THE STANDARD VARIATIONAL PRINCIPLE

There is in quantum mechanics a standard variational principle, called Rayleigh-Ritz method, which is this: If H is the hamiltonian of the system, whose lowest energy value is E_0, then with f representing any arbitrary function

$$E_0 \leq \int f^* H f\, d(\text{vol})\, \left[\int f^* f\, d(\text{vol})\right]^{-1} \qquad (11\text{-}33)$$

This has wide application and is very easily demonstrated. If the function f be expanded as a series in the eigenfunction ϕ_n belonging to the hamiltonian, i.e., if $f = \Sigma a_n \phi_n$, it is evident that

$$\int f^* H f\, d(\text{vol}) \left[\int f^* f\, d(\text{vol})\right]^{-1} = \sum_n |a_n|^2 E_n \left[\sum_n |a_n|^2\right]^{-1} \qquad (11\text{-}34)$$

This latter expression is an average of the energy values (with positive weights $|a_n|^2$) which therefore exceeds (or equals) the least value E_0. The principle expressed in Eq. (11-33) has characteristics similar to the principle of Eq. (11-13). In fact, Eq. (11-33) is a special case of Eq. (11-13). (To be more precise, we should restrict this conclusion to those cases for which the hamiltonian H is derived from a lagrangian which does not contain any magnetic field. Under this restriction, then, the conclusion holds.) To see the relation between these two equations, we shall consider the following example:

Suppose the action S is connected with a lagrangian such as

$$L = \tfrac{1}{2}m\dot{x}^2 - V(x) \tag{11-35}$$

where $V(x)$ is independent of t. (Otherwise, of course, there are no fixed energy levels to seek!) We shall limit ourselves to the case of a single variable x, but the general case follows directly. We note here that if the lagrangian contains the term $\dot{x}A$—for example, if the lagrangian represents a particle in a magnetic field—then Eq. (11-33) is still correct. However, the action S is complex. In this case we suspect that Eq. (11-13) (or some simple modification of this equation) is still valid. However, this has not been proved. So, for the present we shall limit our discussion to a case in which no magnetic field is present. Then in the limit for large values of β we have

$$\int \left(\exp\left\{-\int_0^\beta \frac{1}{2}m\dot{x}^2\,dt + \int_0^\beta V[x(t)]\,dt\right\}\right)\mathcal{D}x(t) \sim e^{-\beta E_0} \tag{11-36}$$

Now suppose we use for our trial action S' the form

$$S' = \int_0^\beta \frac{1}{2}m\dot{x}^2\,dt - \int_0^\beta V'[x(t)]\,dt \tag{11-37}$$

which involves some other potential $V'(x)$. This means that

$$S - S' = \int_0^\beta \{V'[x(t)] - V[x(t)]\}\,dt \tag{11-38}$$

or

$$\delta = -\int e^{S'}\frac{1}{\beta}\int_0^\beta \{V[x(t)] - V'[x(t)]\}\,dt\,\mathcal{D}x(t)\left[\int e^{S'}\,\mathcal{D}x(t)\right]^{-1} \tag{11-39}$$

If we were to define the mean value of any function which depends on the path $x(t)$ in such a manner as this, we would find that the value is nearly independent of t so long as t was not too close to either 0 or β. Therefore, to a sufficient approximation, we can write

$$\begin{aligned}\delta &= -\textstyle\int e^{S'}\{V[x(t)] - V'[x(t)]\}\,\mathcal{D}x(t)\,[\textstyle\int e^{S'}\,\mathcal{D}x(t)]^{-1}\\ &= \langle V[x(t)] - V'[x(t)]\rangle\end{aligned} \tag{11-40}$$

Following the methods given in Chap. 2, we can evaluate this path integral if we assume that the functions ϕ'_n and E'_n belonging to the S'

are known. If our path goes from x_1 to x_2, for example, we obtain

$$\langle f[x(t)]\rangle = \sum_n \{\exp[-(\beta - t)E'_m]\}[\exp(-tE'_n)]\phi'_n(x_2)\phi'_m(x_1)f_{nm} \times \Big\{\sum_n [\exp(-\beta E'_n)]\phi'^*_n(x_2)\phi'_n(x_1)\Big\}^{-1} \tag{11-41}$$

where

$$f_{nm} = \int \phi'^*_n(x) f(x) \phi'_m(x)\, dx \tag{11-42}$$

But, if β approaches infinity and t is likewise large (for example, $t = \beta/2$), all the higher exponentials are negligible compared to the exponential involving the lowest-energy term E'_0. Thus in the limit

$$\lim_{\beta \to \infty} \langle f\rangle = f_{00} \tag{11-43}$$

This result can be written as

$$\delta = -\int \phi'^*_0 V(x) \phi'_0\, dx + \int \phi'^*_0 V'(x) \phi'_0\, dx \tag{11-44}$$

Of course, we must subtract this value from E'_0. However, if H' is the hamiltonian associated with S', that is, if

$$H' = \frac{p^2}{2m} + V'(x) \tag{11-45}$$

then

$$H'\phi'_0 = E'_0\phi'_0 \tag{11-46}$$

so that

$$E'_0 - \delta = \int \phi'^*_0 H' \phi'_0\, dx + \int \phi'^*_0 V \phi'_0\, dx - \int \phi'^*_0 V' \phi'_0\, dx \tag{11-47}$$

But the true hamiltonian can be written as

$$H = \frac{p^2}{2m} + V = \frac{p^2}{2m} + V' + V - V' = H' + V - V' \tag{11-48}$$

and this means that

$$E_0 \leq \int \phi'^*_0 H \phi'_0\, dx \tag{11-49}$$

where ϕ'_0 is normalized and is the wave function corresponding to the lowest energy state of the hamiltonian given in Eq. (11-45). We note that the estimate of the lowest energy level, as given in Eq. (11-49), involves the arbitrary potential $V'(x)$ only through the wave function ϕ'_0. Since this potential was arbitrary, so is the wave function ϕ'_0. Therefore, instead of choosing an arbitrary potential and finding from it the resulting wave function and then proceeding to evaluate

Eq. (11-49), we could instead pick the wave function itself and then evaluate Eq. (11-49) without ever bothering about the potential to which this arbitrary wave function belongs. The variable function in this process is then the wave function ϕ_0' rather than the potential function $V'(x)$. We find, then, that this result is simply another way of stating the result of Eq. (11-33).

If the problems such as the one given in this example were the only ones in which the concept expressed in Eq. (11-13) were useful, then there would not be much point to this long discussion. But there are much more complicated integrals for which Eq. (11-13) can be used in a way that, at least as far as we can tell, is not so easily transformable into Eq. (11-33). We shall describe such an example in the next section.

11-4 SLOW ELECTRONS IN A POLAR CRYSTAL[1]

We imagine an electron moving in a polar crystal, such as sodium chloride. The electron interacts with ions, which are not rigidly fixed. Thus, the electron creates in its neighborhood a distortion of the crystal lattice, and if the electron moves about, the region of distortion moves with it. This electron, together with its distorted environment, has been called a *polaron*.

One consequence of the distortion of the lattice is that the energy of the electron is lowered. Furthermore, since as the electron moves the ions must move to adjust the distortions, the effective inertia of the electron (or to use the currently accepted term, the mass of the polaron) is higher than simply the mass the electron would obtain if the lattice were composed of rigidly fixed points. The precise motion of such a polaron analyzed quantum-mechanically is exceedingly complicated. We shall, however, make a number of approximations whose justification in the real case may be quite difficult. Nevertheless, we shall arrive at an idealized problem which has been studied by a number of physicists.[2] It has been studied not only because of its possible connection with the real behavior of an electron and a crystal, but also because it represents one of the simplest examples of the interaction of a particle and a field. The path integral variational method has been very successful in the solution of this idealized problem.

---

[1] R. P. Feynman, Slow Electrons in a Polar Crystal, *Phys. Rev.*, vol. 97, p. 660, 1955.

[2] For example, H. Frohlich writing in *Advanc. Phys.*, vol. 3, p. 325, 1954. References to other works are given in this article.

First, we note that even if the ions were rigidly fixed in the crystal, the electron would still move in a very complicated potential. In such a case, one can show that there are solutions of the Schrödinger equation for the electron with characteristic wave numbers **k**. The energy levels of these solutions are generally very complicated functions of the wave number. Nevertheless, we assume that the relation between the energy E and the wave number **k** is still a quadratic form, such as

$$E = \frac{\hbar^2 k^2}{2m} \qquad (11\text{-}50)$$

where m is a constant (not necessarily the mass of an electron in a vacuum). Next, we note that the force which the electron exerts on the lattice is such as to push away the negative ions and attract the positive ions. The motion of these ions will be analyzed by considering them as a set of harmonic oscillators and employing the methods of Chap. 8. However, we shall assume that the only harmonic modes which we need are those with high frequency, in which ions of opposite sign of charge move in opposite directions. The frequency $\omega_{\mathbf{k}}$ of each mode then really depends on the wave number **k** of the mode. However, we shall neglect this dependence and assume that ω is a constant.

Our object is to find the electrical force generated by a distortion characterized by the wave number **k** and find the interaction of the electron in this force. Here, we neglect the atomic structure and treat the material of our crystal as simply a continuous dielectric which carries waves of polarization. If **P** is the polarization, written in the form of a longitudinal wave

$$\mathbf{P} = \frac{\mathbf{k}}{k} a_k e^{i\mathbf{k}\cdot\mathbf{r}} \qquad (11\text{-}51)$$

then the charge density from the ions is

$$\rho = \nabla \cdot \mathbf{P} = k a_k e^{i\mathbf{k}\cdot\mathbf{r}} \qquad (11\text{-}52)$$

If the potential is V, we have

$$\nabla^2 V = \rho \qquad (11\text{-}53)$$

Thus if q_k is the amplitude of the kth longitudinal running wave, the polarization a_k is proportional to q_k, and the interaction between the wave of polarization and the electron is proportional to the sum over all values of **k** of the terms $(q_k/k)e^{i\mathbf{k}\cdot\mathbf{x}}$.

Since the energy and the momentum of the electron are related by the expression $E = p^2/2m$, we can write the lagrangian of the entire system as

$$L = \frac{1}{2}|\dot{\mathbf{r}}|^2 + \sum_{\mathbf{k}} \frac{1}{2}(\dot{q}_k^2 - q_k^2) + \left(\frac{2\sqrt{2}\,\pi\alpha}{V}\right)^{1/2} \sum_{\mathbf{k}} \frac{1}{k} q_k e^{i\mathbf{k}\cdot\mathbf{r}} \tag{11-54}$$

The first term of the expression is the energy of the electron in a rigid lattice, where $\mathbf{r}$ is its position. The second term represents the lagrangian of the oscillations of polarizations taken alone, where it is assumed that all waves of polarization have the same frequency and the coordinate of the kth mode is q_k. The last term is the lagrangian of the interaction between the electron and the lattice vibrations, where V in this expression represents the volume of the crystal and α is a constant. To simplify writing of all our subsequent formulas, we have written this in dimensionless form. That is, the scales of energy, length, and time are so chosen that not only $\hbar$ but also the common frequency ω of the oscillators and the mass m of the electron are all unity. The coupling constant α is then the dimensionless ratio

$$\alpha = \frac{1}{\sqrt{2}}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon}\right)e^2 \tag{11-55}$$

where ϵ and ϵ_∞ are the static and high-frequency dielectric constants, respectively. In a typical case, such as the crystal of NaCl, the value of α may be about 5. The values of energy which we shall calculate are in units of $\hbar\omega$.

Now we can study the quantum-mechanical motion of the electron, solving the motion of the harmonic oscillators completely. For example, the amplitude that the electron starts at $\mathbf{r}_1$ with the oscillators in the ground state and ends at the position $\mathbf{r}_2$ with the oscillators still in the ground state is

$$G_{00}(2,1) = \int e^{iS}\, \mathcal{D}\mathbf{r}(t) \tag{11-56}$$

where we have used the results of Chap. 8, and

$$S = \frac{1}{2}\int \left|\frac{d\mathbf{r}}{dt}\right|^2 dt + \int \frac{\sqrt{2}\,\pi\alpha}{k^2} e^{i\mathbf{k}\cdot\mathbf{r}(t)} e^{-i\mathbf{k}\cdot\mathbf{r}(s)} e^{-i|t-s|}\, dt\, ds\, \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{11-57}$$

If we complete the integral over the wave numbers $\mathbf{k}$, we find

$$S = \frac{1}{2}\int |\dot{\mathbf{r}}|^2\, dt + \frac{\alpha i}{\sqrt{8}} \int \frac{e^{-i|t-s|}}{|\mathbf{r}(t) - \mathbf{r}(s)|}\, dt\, ds \tag{11-58}$$

This quantity $G_{00}(2,1)$ will depend upon the initial and final positions of the electron, $\mathbf{r}_1$ and $\mathbf{r}_2$, and upon the time interval we are considering,

T. Since this function is a kernel, it is a solution of the Schrödinger wave equation, considered as a function of the time interval T. Therefore, we realize that it will contain frequencies in its exponentials which are proportional to the energy levels E_m. It is the lowest one of these energy levels which we now seek.

In developing our variational principle, as we have explained, we are not interested in the kernel for real time intervals T. Instead, we want quantities such as those which appear in Eq. (11-8) for large values of β. By following all the steps leading to Eq. (11-58), it can be readily shown for imaginary values of the time variable β that the resulting kernel has the form

$$K(2,1) = \int e^{S}\,\mathcal{D}\mathbf{r}(t) \tag{11-59}$$

where the variable t goes from 0 to β and

$$S = \frac{1}{2}\int \left|\frac{d\mathbf{r}}{dt}\right|^2 dt + \frac{\alpha}{\sqrt{8}}\iint \frac{e^{-|t-s|}}{|\mathbf{r}(t)-\mathbf{r}(s)|}\,dt\,ds \tag{11-60}$$

This result is just that which one might expect from the replacement of t by the imaginary value it in Eq. (11-58). Asymptotically, for large values of β, this kernel becomes proportional to $e^{-E_0\beta}$.

We now have a relatively complicated path integral on which to try our variational principle. Next, we shall have to choose some simple action S', which roughly approximates the true action S, and then find E' and δ.

We note that in Eq. (11-60) the particle considered at any particular time[1] "interacts" with its position at a past time by a reaction which is inversely proportional to the distance traveled between these two times and which dies out exponentially with the time difference. The reason for this is that the disturbance set up by the electron in the crystal lattice in the past takes some time to die out. That is, it takes some time for the ions to relax, and here in this relaxation period the electron still "feels" the old disturbance.

We shall try an action S' which has this same property, except that instead of involving the inverse distance as a coupling law, the attraction will have the geometric form of the parabolic well. This would be a poor approximation if the distance $|\mathbf{r}(s) - \mathbf{r}(t)|$ could very often become exceedingly large. However, since there is a limited time available before the exponential time factor forces the interaction to die out, large values of this difference will not make any important

[1] Although t in Eq. (11-60) is not really the time, but an integration variable instead, it is useful to think about it as a time, just as we did in Sec. 10-2.

contribution to the integral. Thus, we shall try

$$S' = -\tfrac{1}{2}\int|\dot{\mathbf{r}}|^2\,dt - \tfrac{1}{2}C\int\int|\mathbf{r}(t) - \mathbf{r}(s)|^2 e^{-w|t-s|}\,dt\,ds \tag{11-61}$$

The constant C is a measure of the strength of the attraction between the electron and the previously created disturbance. We take this as an adjustable parameter. Furthermore, we can with no extra difficulty permit the exponential cutoff law to contain the adjustable constant w, which can be different from unity. With this extra parameter we can partly compensate for the imperfection which we have introduced by replacing the inverted distance effect by a parabolic effect. [We also note in this regard that adding an extra constant to the parabolic term $|\mathbf{r}(t) - \mathbf{r}(s)|^2$ leads to no further freedom, since such a term would drop out in evaluating a formula for E_0'.] We shall adjust variable parameters C and w later in the evaluation in order to make E_0' a minimum.

Since the action S' we have picked is quadratic, all of the path integrals which result are easily worked out by the methods described in Chap. 2.

By comparing Eqs. (11-60) and (11-61), we find that

$$\begin{aligned}\frac{1}{\beta}\langle S - S'\rangle &= \frac{\alpha}{\sqrt{8}}\int\left\langle\frac{1}{|\mathbf{r}(t) - \mathbf{r}(s)|}\right\rangle e^{-|t-s|}\,ds \\ &\qquad + \frac{1}{2}C\int\langle|\mathbf{r}(t) - \mathbf{r}(s)|^2\rangle e^{-w|t-s|}\,ds \\ &= A + B\end{aligned} \tag{11-62}$$

We shall concentrate our attention on the first term on the right-hand side of this equation, A. In this term we can express $|\mathbf{r}(t) - \mathbf{r}(s)|^{-1}$ by a Fourier transform. As a matter of fact, this term is the result of the Fourier transform involved in the step between Eqs. (11-57) and (11-58). So we have

$$|\mathbf{r}(t) - \mathbf{r}(s)|^{-1} = \int d^3\mathbf{k}\,\exp\{i\mathbf{k}\cdot[\mathbf{r}(t) - \mathbf{r}(s)](2\pi^2k^2)^{-1}\} \tag{11-63}$$

For this reason we need to study

$$\langle\exp\{i\mathbf{k}\cdot[\mathbf{r}(\tau) - \mathbf{r}(\sigma)]\}\rangle = \frac{\int(e^{S'}\exp\{i\mathbf{k}\cdot[\mathbf{r}(\tau) - \mathbf{r}(\sigma)]\})\,\mathcal{D}\mathbf{r}(t)}{[\int e^{S'}\,\mathcal{D}\mathbf{r}(t)]^{-1}} \tag{11-64}$$

The integral in the numerator is of the form

$$\begin{aligned}I = \int\Bigg\{\exp\Bigg[-\frac{1}{2}\int\left|\frac{d\mathbf{r}}{dt}\right|^2 dt &- \frac{1}{2}C\iint|\mathbf{r}(t) - \mathbf{r}(s)|^2 e^{-w|t-s|}\,dt\,ds \\ &+ \int\mathbf{f}(t)\cdot\mathbf{r}(t)\,dt\Bigg]\Bigg\}\,\mathcal{D}\mathbf{r}(t)\end{aligned} \tag{11-65}$$

where specifically

$$\mathbf{f}(t) = i\mathbf{k}\delta(t - \tau) - i\mathbf{k}\delta(t - \sigma) \tag{11-66}$$

Now we shall evaluate Eq. (11-65) in so far as it depends on **f** or **k** aside from a normalization factor which drops out in Eq. (11-64). Incidentally, let us notice that the three rectangular components separate in Eq. (11-65) and we need consider only a scalar case. The method of integration is the same as that introduced in Chap. 2 for the evaluation of gaussian path integrals. Thus we substitute $X(t) = X'(t) + Y(t)$, where $X'(t)$ is that special function for which the exponent is maximum. The variable of integration is now $Y(t)$. Since the exponent is quadratic in $X(t)$ and X' renders it an extremum, it can contain $Y(t)$ only quadratically; so Y then separates off as a factor not containing f, which may be integrated to give an unimportant constant (depending on T only). Therefore, within such a constant

$$I = \exp\left[-\tfrac{1}{2}\int \dot{X}'^2(t)\,dt - \tfrac{1}{2}C\int\int [X'(t) - X'(s)]^2 e^{-w|t-s|}\,dt\,ds + \int f(t)X'(t)\,dt\right] \tag{11-67}$$

where X' is that function which minimizes the expression [subject for convenience to the boundary condition $X'(0) = X'(T) = 0$ if the time interval is 0 to T]. The variational problem gives the integral equation

$$\frac{d^2X'(t)}{dt^2} = 2C\int [X'(t) - X'(s)]e^{-w|t-s|}\,ds - f(t) \tag{11-68}$$

Using Eq. (11-68), Eq. (11-67) can be simplified to

$$I = \exp\left[\int \tfrac{1}{2} f(t)X'(t)\,dt\right] \tag{11-69}$$

We need merely solve Eq. (11-68) and substitute into Eq. (11-69). To do this, we define

$$Z(t) = \frac{w}{2}\int e^{-w|t-s|}\,X'(s)\,ds \tag{11-70}$$

so that

$$\frac{d^2Z(t)}{dt^2} = w^2[Z(t) - X'(t)] \tag{11-71}$$

while Eq. (11-68) is

$$\frac{d^2X'(t)}{dt^2} = \frac{4C}{w}[X'(t) - Z(t)] - f(t) \tag{11-72}$$

The equations are readily separated and solved. The solution for $X'(t)$ substituted into Eq. (11-69) gives, for the case of Eq. (11-68),

$$I = \exp\{i\mathbf{k}\cdot[\mathbf{X}(\tau) - \mathbf{X}(\sigma)]\}$$

$$= \exp\left[-\frac{2Ck^2}{v^3w}(1 - e^{-v|\tau-\sigma|}) - \frac{w^2}{2v^2}k^2|\tau - \sigma|\right] \tag{11-73}$$

where we have made the substitution

$$v^2 = w^2 + \frac{4C}{w} \tag{11-74}$$

The result is correctly normalized, since it is valid for $\mathbf{k} = 0$. Upon substitution from Eq. (11-73) into Eq. (11-63) there results an integral on $\mathbf{k}$ which is a simple gaussian, so that substitution into A gives

$$A = \pi^{-1/2}\frac{\alpha v}{w}\int_0^\infty \left[w^2\tau + \frac{v^2 - w^2}{v}(1 - e^{-v\tau})\right]^{-1/2} e^{-w\tau}\,d\tau \tag{11-75}$$

To find B, we need $\langle|\mathbf{r}(t) - \mathbf{r}(s)|^2\rangle$. This can be obtained by expanding both sides of Eq. (11-73) with respect to $\mathbf{k}$ up to order k^2. Therefore

$$\frac{1}{3}\langle|\mathbf{r}(\tau) - \mathbf{r}(\sigma)|^2\rangle = \frac{4C}{v^3w}(1 - e^{-\tau-\sigma}) + \frac{w^2}{v^2}|\tau - \sigma| \tag{11-76}$$

The integral in B is now easily performed and the expression simplifies to

$$B = \frac{3C}{vw} \tag{11-77}$$

Finally we need E', the energy belonging to our action S'. This is most easily obtained by differentiating both sides of Eq. (11-6) with respect to C. One finds immediately

$$\frac{C\,dE_0'}{dC} = B \tag{11-78}$$

so that, in view of Eqs. (11-77) and (11-74), integration gives

$$E_0' = \tfrac{3}{2}(v - w) \tag{11-79}$$

since $E_0' = 0$ for $C = 0$. Since $E_0' - B = (3/4v)(v - w)^2$, we finally obtain for our energy expression

$$E = \frac{3}{4v}(v - w)^2 - A \tag{11-80}$$

with A given in Eq. (11-75). The quantities v and w can be considered as two parameters which may be varied separately to obtain a minimum.

The integral in A, unfortunately, cannot be performed in closed form, so that a complete determination of E requires numerical integration. It is, however, possible to obtain approximate expressions in various limiting cases. The case of large α corresponds to large v. The choice $w = 0$ leads to an integral

$$A = \pi^{-1/2}\alpha v^{1/2} \int_0^\infty e^{-\tau}\, d\tau\, (1 - e^{-v\tau})^{-1/2} = \frac{\alpha\Gamma(1/v)}{v^{1/2}\Gamma(1/2 + 1/v)} \tag{11-81}$$

and $E_0' = 3v/4$. It corresponds to the use of a fixed harmonic binding potential in Eq. (11-37). For large v, $e^{-v\tau}$ can be neglected, so that $A = (\pi\omega)^{-1/2}\alpha v^{1/2}$. For α less than 5.8 and $w = 0$, Eq. (11-80) does not give a minimum unless $v = 0$, so that the $w = 0$ case does not give a single expression for all ranges of α. In spite of this disadvantage, the result with Eq. (11-81) is relatively simple and fairly accurate. For $\alpha > 6$, only fairly large values of v are important, and the asymptotic formula (good to 1 per cent for $v > 4$)

$$A = \alpha\left(\frac{v}{\pi}\right)^{1/2}\left(1 + \frac{2\ln 2}{v}\right) \tag{11-82}$$

is convenient. Frohlich, however, considers the discontinuity at $\alpha = 6$ as a serious disadvantage—a disadvantage which can be avoided in our present approach. This we do by choosing w different from zero.

Let us study Eq. (11-80) just for small α, in case w is not zero. The minimum will occur for v near w. Therefore, we write $v = (1 + \epsilon)w$, consider ϵ small, and expand the root in Eq. (11-80). This gives

$$A = \alpha\frac{v}{w}\left[1 - \epsilon\int_0^\infty \tau^{-3/2}e^{-\tau}(1 - e^{w\tau})\,\frac{d\tau}{2\pi^{1/2}} + \cdots\right] \tag{11-83}$$

The integral is

$$2w^{-1}[(1 + w)^{1/2} - 1] = P \tag{11-84}$$

The problem of Eq. (11-80) then corresponds, in this order, to minimizing

$$E = \tfrac{3}{4}w\epsilon^2 - \alpha - \alpha\epsilon(1 - P) \tag{11-85}$$

That is,

$$\epsilon = \frac{2\alpha(1-P)}{3w} \tag{11-86}$$

which is valid for small α only, because ϵ was assumed small. The resulting energy is

$$E = -\alpha - \frac{\alpha^2(1-P)^2}{3w} \tag{11-87}$$

Our method therefore gives a correction even for small α. It is least for $w = 3$, in which case it gives

$$E = -\alpha - \frac{\alpha^2}{81} = -\alpha - 1.23\left(\frac{\alpha}{10}\right)^2 \tag{11-88}$$

It is not sensitive to the choice of w. For example, for $w = 1$ the 1.23 falls only to 0.98. The method of Lee and Pines[1] gives exactly the result of Eq. (11-88) to this order. The perturbation expansion has been carried to second order by Haga,[2] who shows that the exact coefficient of the $(\alpha/10)^2$ term should be 1.26, so that our variational method is remarkably accurate for small α.

The opposite extreme of a large α corresponds to large v and, as we shall see, w near 1. Since $v \gg w$, the integral in Eq. (11-76) reduces in the first approximation to Eq. (11-81), which we can use in its asymptotic form. The next approximation in w can be obtained by expanding the radical in Eq. (11-75), considering $w/v \ll 1$. Furthermore, $e^{-v\tau}$ is negligible. In this way we get

$$E = \frac{3}{4v}(v-w)^2 - \alpha\left(\frac{v}{\pi}\right)^{1/2}\left(1 + \frac{2\ln 2}{v} - \frac{w^2}{2v}\right) \tag{11-89}$$

This is minimum, within our approximation of large v, when $w = 1$ and $v = (4\alpha^2/9\pi) - (41n^2 - 1)$. Then we find[3]

$$E = -\frac{\alpha}{3\pi} - 3\ln 2 - \frac{3}{4} = -0.1061\alpha^2 - 2.83 \tag{11-90}$$

---

[1] T. Lee and D. Pines, Interaction of a Nonrelativistic Particle with a Scalar Field with Application to Slow Electrons in Polar Crystals, *Phys. Rev.*, vol. 92, p. 883, 1953.

[2] E. Haga, Note on the Slow Electrons in a Polar Crystal, *Progr. Theoret. Phys. (Kyoto)*, vol. 11, p. 449, 1954.

[3] S. I. Pekar in Theory of Polarons, *Zh. Eksperim. i Teor. Fiz.*, vol. 19, p. 796, 1949, has shown that E goes as $-0.1088\alpha^2$ for the case of large α.

TABLE 11-1

α	3	5	7	9	11
v	3.44	4.02	5.81	9.85	15.5
w	2.55	2.13	1.60	1.28	1.15
E_f	−3.1333	−5.4401	−8.1127	−11.486	−15.710
E_{1p}	−3.10	−5.30	−7.58	− 9.95	−12.41
E_g	−3.09	−5.24	−7.43	− 9.65	−11.88
E_{pbt}			−6.83	−10.31	−14.7

The approximations do not keep E as an upper limit because, unfortunately, the further terms, of order $1/\alpha^2$, are probably positive.

Detailed and numerical work based on this approach has been carried out by T. D. Schultz.[1] Using a digital computer, Schultz worked out values of v and w which would give a minimum for several different values of α. He also evaluated E and compared it with the values which would be obtained from several alternative theories. In particular, he worked out the self-energy from the theories of Lee, Low, and Pines (E_{11p}),[2] Lee and Pines (E_{1p}),[3] Gross (E_g),[4] and Pekar,[5] Bogolubov,[6] and Tyablikov[7] (E_{pbt}).

The results for α, e, and w and also for the energies given by the Feynman theory (E_f) compared with energies derived from the other theories are given in Table 11-1, which is reproduced from the paper of Schultz. In this table, both ω and $\hbar$ are assumed to have the value 1. Note that for all values of α, the value of E_f is less than all others.

---

[1] T. D. Schultz, Slow Electrons and Polar Crystals: Self-energy, Mass, and Mobility, *Phys. Rev.*, vol. 116, p. 526, 1959.

[2] Lee, Low, and Pines, The Motion of Slow Electrons in a Polar Crystal, *Phys. Rev.*, vol. 90, p. 297, 1953.

[3] *Op. cit.*

[4] E. P. Gross, Small Oscillation Theory of the Interaction of a Particle and a Scalar Field, *Phys. Rev.*, vol. 100, p. 1571, 1955.

[5] S. I. Pekar, "Untersuchungen über die Elektronentheorie der Kristalle," Akademie-Verlag, Berlin, 1954.

[6] S. I. Bogolubov, On a New Form of the Adiabatic Theory of Disturbances in the Problem of Interaction of Particles with a Quantum Field, *Ukr. Mat. Zh.*, vol. 2, p. 3, 1950.

[7] S. V. Tyablikov, An Adiabatic Form of Perturbation Theory in the Problem of the Interaction of a Particle with a Quantum Field, *Zh. Exsperim. i Teor. Fiz.*, vol. 21, p. 377, 1951.

12

Other Problems in Probability

In the preceding chapters we have seen how to use path integrals to treat a number of quantum-mechanical problems which are, by their very physical nature, probabilistic problems. We have also used the path integral method to analyze some aspects of statistical mechanics wherein the probabilistic nature of the functions permitted the path integral technique to be particularly effective. We can continue this line of development into a wide variety of probability problems where this approach has special and valuable applications.

It is the purpose of this chapter to explore a number of these probability problems. They will be of two kinds. First we shall discuss direct applications of path integral ideas to classical probability problems. This is quite different from all preceding chapters, in which all applications were to quantum mechanics. Following that, we shall deal with problems in which both probability and quantum mechanics are involved. We cannot, in this chapter, deal with these matters in any detail. We shall only outline by some examples how certain problems may be set up and thereby suggest to the reader other applications of the path integral approach.

The main direct application of path integrals to probability problems is due to the ability of path integrals to deal directly with the notion of the probability of a path or function. To make this idea clear, we proceed in steps from the well-known[1] ideas of probability applied to discrete events and to continuous variables.

12-1 RANDOM PULSES

To start with, suppose we consider a typical probability problem for a discrete variable. We are given a situation in which a series of discrete events is taking place at random times, e.g., cosmic rays striking a detector or raindrops falling on a specifically demarked area of ground. We know that the particles fall at random times, but in any long period of time T we expect $\bar{n} = T\mu$ particles will be observed. In other words, μ is the mean counting rate.

Of course, in any actual measurement the exact number of particles n recorded will not, in general, correspond to the expected number. But we can ask directly: What is the probability of observing a particular number n of particles during a period when the expected number of particles is $\bar{n}$? It is given by the Poisson distribution

$$P_n = \frac{\bar{n}^n}{n!e^{\bar{n}}} \tag{12-1}$$

---

[1] Harold Cramér, "Mathematical Methods of Statistics," Princeton University Press, Princeton, N.J., 1951. We assume knowledge of usual probability theory.

On the other hand, we might ask a probability question of a different kind. We might, for example, ask: What is the probability that the interval from one particle impact to the next will be some particular time t? Actually, there is no correct answer to the question phrased this way. If we were to ask the probability that the time interval will be equal to or greater than t, then we could give an answer [it is $\exp(-\mu t)$]. That is, we can get an answer to a question about t falling within a certain range. Thus, if we are interested in a particular value, we must allow ourselves an infinitesimal range and ask the question: What is the (infinitesimal) probability that the time interval will fall within the range dt centered around t? The answer is written as

$$P(t)\,dt = \mu e^{-\mu t}\,dt \tag{12-2}$$

So we create a concept of a probability distribution of a continuous variable: $P(t)$ is the probability per unit range of t that the interval is t. We write the probability distribution of x as $P(x)$ if $P(x)\,dx$ is the probability that the variable lies in the range dx about x. We can easily extend this to two variables and write the probability distribution of x and y as $P(x,y)\,dx\,dy$. By this we mean that the probability of finding the variables x and y in the region R of the xy plane is given by $\int_R P(x,y)\,dx\,dy$.

We wish to expand the concepts of probability still further. We want to consider the distribution not of single variables but of complete curves; i.e., we want to construct probability functions, or rather functionals, which will permit us to answer the question: What is the probability of obtaining a particular time history of a physical phenomenon, such as the voltage in a resistor or the price of a commodity, or, in two variables, the probability of a certain shape of the surface of the sea as a function of latitude and longitude? Thus, we are led to consider the *probability of a function*.

We shall write it down this way. The probability of observing the function $f(t)$ is a functional $P[f(t)]$. But we must be careful to remember that questions relating to such a probability have meaning only if we define the range within which we are looking for a specific curve. Just as in the example above we had to ask the question: What is the probability of finding the time interval within the range dt? so now we must ask: What is the probability of finding the function within some more or less restricted class of function, for example, those curves which are bounded between values a and b for the complete time history in which we are interested? If we call such a subset of curves the class A, then we ask: What is the probability of finding $f(t)$ in the class

A? and we write the answer as the path integral

$$\int_A P[f(t)]\,\mathfrak{D}f(t) \tag{12-3}$$

where the integral extends over all functions of class A.

Actually, this expression can be thought of as similar to the probability function for a number of different variables. If we imagine time to be divided into discrete intervals (as we imagined it when we were first defining path integrals in Chap. 2) taking on the values of t_1, t_2, . . . , then the values of the function at those particular times $f(t_1)$, $f(t_2)$, . . . $= f_1, f_2$, . . . are analogous to the variables of a multivariable distribution function. The probability of observing a particular curve can then be thought of as the probability of obtaining a particular set of values f_1, f_2, . . . in the range df_1, df_2, . . . , that is, $P(f_1, f_2, \ldots)\, df_1\, df_2 \cdots$.

If we then proceed to the limit as the number of discrete intervals in time becomes infinite, we obtain the probability of observing the continuous curve $f(t)$ in the range $\mathfrak{D}f(t)$ as the integrand in the path integral of Eq. (12-3). It is this probability concept and this probability functional with which we shall be working in the remainder of this chapter.

12-2 CHARACTERISTIC FUNCTIONS

It is helpful to continue using the analogy between the probability functional of a path and the more traditional probability function of a variable. A number of concepts, such as the concept of a mean value, are common to the two approaches. With usual probability distributions for quantities which have discrete values, so that the probability of observing the specific number n is P_n, the mean is

$$\sum_{n=1}^{\infty} nP_n = \bar{n} \tag{12-4}$$

For a continuously distributed variable, it is

$$\int_{-\infty}^{\infty} xP(x)\,dx = \bar{x} \tag{12-5}$$

and in an analogous fashion, the mean value of the functional $Q[f(t)]$ is written

$$\frac{\int Q[f(t)]P[f(t)]\,\mathfrak{D}f(t)}{\int P[f(t)]\,\mathfrak{D}f(t)} = \langle Q\rangle \tag{12-6}$$

In this last equation, as in Chap. 7, we have included a path integral in the denominator to remind ourselves that we are always faced with normalizing problems. In principle, it would be possible to work out the path integral of the distribution function, set it equal to 1, and so evaluate the normalizing constant to begin with. However, in many practical cases it is more convenient to leave the function unnormalized and simply cancel out factors on the top and bottom of the expression which might, in actuality, be extremely difficult to evaluate.

Just as the mean value of the function can be expressed in the path integral notation, so can the mean-square value of the function at a particular time, say, $t = a$. Thus,

$$\langle [f(a)]^2 \rangle = \frac{\int [f(a)]^2 P[f(t)]\, \mathfrak{D}f(t)}{\int P[f(t)]\, \mathfrak{D}f(t)} \tag{12-7}$$

for this is only a special functional.

One of the most important mean values of a function, as evaluated with Eq. (12-5), is the mean of e^{ikx}. It is called the characteristic function, and it is

$$\phi(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} e^{ikx} P(x)\, dx \tag{12-8}$$

This is sometimes also called the moment-generating function. It is simply the Fourier transform of $P(x)$, and it is an extremely useful function for evaluating various characteristics of the distribution, since it is equivalent to a knowledge of the distribution function itself. This last fact is the result of the possibility of performing the inverse transform as

$$P(x) = \int_{-\infty}^{\infty} e^{-ikx} \phi(k)\, dk \tag{12-9}$$

A number of important parameters of the distribution can be determined by taking the derivatives of the characteristic function. Thus, for example, the mean value of x is

$$\langle x \rangle = -i \frac{d\phi(k)}{dk}\bigg|_{k=0} \tag{12-10}$$

as is readily demonstrated by differentiating each expression in Eq. (12-8) with respect to k and then setting $k = 0$. In fact, a series of such relations exists:

$$\phi(0) = 1 \qquad \phi'(0) = i\langle x \rangle \qquad \phi''(0) = -\langle x^2 \rangle \qquad \cdots \tag{12-11}$$

Of course, our next step is to generalize the concept of the characteristic function to the functional distribution case. We could construct a mathematical definition of such a characteristic function by returning again to our picture of discrete time intervals. We then wish to perform the Fourier transform on the probability function of a large number of variables, using the kernel $\exp(ik_1f_1)\exp(ik_2f_2)\cdots$. As we go to the limit of an infinite number of time intervals, this becomes simply $e^{i\int k(t)f(t)\,dt}$. This, then, is the functional whose mean value we wish to take in order to develop the *characteristic functional*. By using Eq. (12-6), we obtain

$$\Phi[k(t)] = \frac{\int e^{i\int k(t)f(t)\,dt}P[f(t)]\,\mathfrak{D}f(t)}{\int P[f(t)]\,\mathfrak{D}f(t)} \tag{12-12}$$

This characteristic functional also has important special properties. For example, $\Phi(0) = 1$, and the mean value of the function $f(t)$ evaluated at the particular time $t = a$ is

$$\langle f(a)\rangle = -i\frac{\delta}{\delta k(a)}\,\Phi[k(t)]\bigg|_{k(t)=0} \tag{12-13}$$

where we have used the technique of the functional derivative as described in Sec. 7-2.

In principle, we can invert our path integral Fourier transform and write the probability functional as

$$P[f(t)] = \int e^{-i\int k(t)f(t)\,dt}\Phi[k(t)]\,\mathfrak{D}k(t) \tag{12-14}$$

where now, of course, the path integral is carried out in the space of the k functions.

We may remark, for use in interpretation later on, that if the function $f(t)$ is not uncertain but is definitely known to be some special function $F(t)$, that is, $P[f(t)]$ is zero for all $f(t)$ except $f(t) = F(t)$, then the characteristic function is

$$\Phi = e^{i\int k(t)F(t)\,dt} \tag{12-15}$$

12-3 NOISE

Suppose we apply the ideas so far developed to a particular example and in the process develop a few more concepts. Let us consider the situation in which we are counting some sort of pulses, perhaps pulses generated by the impact of cosmic rays on a Geiger counter or perhaps thermal-noise pulses in a resistor. In such cases the pulses are not simply discrete spikes of energy but are represented by a rising and falling voltage. Thus, careful inspection of the actual voltage history associated with such a pulse would show that it has the form $g(t)$ for a pulse occurring at $t = 0$. So, if the pulse occurred at t_0, the shape of the voltage curve would be $g(t - t_0)$.

Now, suppose we conduct our counting experiment for the time interval of length T during which a number of pulses centered on the times $t_1, t_2, \ldots, t_n$ occurred. The complete voltage history over this experiment would be $\sum_{j=1}^{n} g(t - t_j)$. Since we know when all the events occurred, our probability function would simply be the representation of certainty, and by use of Eq. (12-15) the corresponding characteristic function becomes

$$\Phi = \exp\left[i \sum_{j=1}^{n} \int k(t) g(t - t_j)\, dt\right] \tag{12-16}$$

But now suppose that we wish to determine the probability of finding a particular time history for the voltage before conducting the experiment. In that case we permit the n events to be randomly distributed with uniform probability over the complete time interval. That is, the probability of an event happening within the time interval dt is dt/T. In this case the characteristic functional becomes

$$\begin{aligned}\Phi &= \int_0^T \exp\left[i \sum_{j=1}^{n} \int k(t) g(t - t_j)\, dt\right] \frac{dt_1}{T} \frac{dt_2}{T} \cdots \frac{dt_n}{T} \\ &= \left(\int_0^T e^{i\int k(t+s) g(t)\, dt}\, \frac{ds}{T}\right)^n \end{aligned} \tag{12-17}$$

We call the expression in parentheses A and write this result as A^n.

If the number of events in the time interval is distributed in such a way that the Poisson distribution applies, i.e., the occurrence of each event is independent of the time of occurrence of any other event and there is a constant rate μ for the expected number of events per unit time, then the expected number of events in the time interval T is $\mu T = \bar{n}$. The characteristic function is

$$\Phi = \sum_n A^n \frac{\bar{n}^n}{n!} e^{-\bar{n}} \tag{12-18}$$

The sum on the right-hand side of this equation is the expansion of an exponential function, so that we can write the characteristic function as

$$\begin{aligned}\Phi = e^{-(1-A)\bar{n}} &= \exp\left[-\mu T \left(1 - \int_0^T e^{i\int k(t+s) g(t)\, dt}\, \frac{ds}{T}\right)\right] \\ &= \exp\left[-\mu \int_0^T \left(1 - e^{i\int k(t+s) g(t)\, dt}\right) ds\right] \end{aligned} \tag{12-19}$$

Thus, we may determine the characteristic function for many different situations. We next go on to discuss this result under various approximate circumstances.

Suppose we imagine that the pulses get very weak while the expected number of pulses per unit time, that is, μ, becomes large. In that case $g(t)$ is small, so we can expand $e^{i\int k(t+s)g(t)\,dt}$ in a power series and we can approximate the characteristic function as

$$\exp\left[i\mu\int_0^T\int_0^T k(t+s)g(t)\,dt\,ds\right] = \exp\left[i\mu G\int_0^T k(t)\,dt\right] \tag{12-20}$$

where we have used the substitution $G = \int g(t)\,dt$, the area of the pulse. This means that Φ is in the form of Eq. (12-15) with $F(t) = \mu G$ (a constant independent of t). That is equivalent to saying that $f(t)$ is certainly μG or, in other words, that there is unit probability for observing the function $f(t) = \mu G$ and zero probability for observing any other $f(t)$. That is to say, the pile-up of a large number of small pulses generates a nearly steady direct voltage of value equal to the number of pulses per second μ times the average voltage G supplied by each. Next, we go to one higher approximation and study the fluctuations or irregularities of this nearly direct voltage.

Equation (12-20) is a first-order approximation to the exponential $e^{i\int k(t+s)g(t)\,dt}$ in the description of the characteristic functional of Eq. (12-19). Suppose now that we go on to the next-order approximation and include the second-order term. This is

$$-\frac{\mu}{2}\iint k(t)g(t+s)\,dt\int k(t')g(t'+s)\,dt'\,ds \tag{12-21}$$

To simplify this expression, we define a function which measures the overlap between two nearby pulses as

$$\lambda(\tau) = \int g(t)g(t+\tau)\,dt \tag{12-22}$$

By use of this substitution, the second-order term is reduced to

$$-\frac{\mu}{2}\int_0^T\int_0^T k(t)k(t')\lambda(t-t')\,dt\,dt' \tag{12-23}$$

Including both first- and second-order terms, the characteristic functional is

$$\Phi = e^{i\mu G\int k(t)\,dt}e^{-(\mu/2)\iint k(t)k(t')\lambda(t-t')\,dt\,dt'} \tag{12-24}$$

The first factor in this expression is the constant average level, which we might call the d-c level if we are thinking about voltage pulses.

We can, if we wish, neglect this level and concentrate only on the variations around it by shifting the origin of $f(t)$. That is, we can always take out a factor $e^{i\int k(t)F(t)\,dt}$ by shifting the origin of $f(t)$ [i.e., by writing $f(t) = F(t) + f'(t)$ and studying the probability distribution of $f'(t)$ and its characteristic functional]. If we make this change of origin, we are in a position to study the fluctuations of voltage around the d-c level.

We note one special approximation to Eq. (12-24) which is often adequate. Generally, $\lambda(\tau)$ is a narrow function of τ. The pulse shape $g(t)$ rises and falls with a finite width, so if two pulses are spaced a very great distance apart, their overlapping area vanishes. This is another way of saying that $\lambda(\tau)$ approaches 0 rapidly as τ becomes large. As a result of this, if $\lambda(\tau)$ is narrow enough, the second term in Eq. (12-24) can be approximated by

$$e^{-(q/2)\int[k(t)]^2\,dt} \tag{12-25}$$

where we have used the substitution $q = \mu \int_{-\infty}^{\infty} \lambda\,d\tau$. This is equivalent to the probability distribution

$$P[f(t)] = e^{-(1/2\sigma)\int[f(t)]^2\,dt} \tag{12-26}$$

Such fluctuations as we are describing here are often called *gaussian noise.*

Characteristics of distribution functionals describing noise functions have been studied extensively in recent years in the theory of communications. A number of characteristics of noise spectra have been defined and evaluated, and we shall carry through similar discussions here and in the next section, where we treat gaussian noise.

Now we shall continue to show, by giving one further example, how characteristic functionals are set up. We shall consider pulses which come at random times all with a given characteristic shape, say, $u(t)$, but each with a different scale height, so a typical pulse is written $au(t)$. We might allow the height a to be either plus or minus. So now we suppose the timings of the pulses are randomly spaced instants t_j and the heights take on random positive and negative values a_j. The resulting function is

$$f(t) = \sum_j a_j u(t - t_j) \tag{12-27}$$

If first we set aside the random nature of the events, we obtain a characteristic functional equivalent to that of Eq. (12-16) as

$$\Phi = \exp\left[i\sum_j a_j \int k(t)u(t - t_j)\,dt\right] \tag{12-28}$$

Next, if we include the presumed random nature of the scale heights of the pulses and say that the probability of obtaining a particular scale height a_j for the jth pulse in the region da_j is $p(a_j)\,da_j$, then the characteristic functional becomes

$$\Phi = \iint \cdots \exp\left[i\sum_j a_j \int k(t)u(t - t_j)\,dt\right] p(a_1)\,da_1\,p(a_2)\,da_2 \cdots \tag{12-29}$$

Of course, each of these probability functions for the values of a_j has associated with it a characteristic function or moment-generating function. Suppose we call this function $W[\omega]$ and define it as

$$W[\omega] = \int e^{i\omega a}p(a)\,da \tag{12-30}$$

Then the expression for Φ can be written as

$$\Phi = \prod_j W\left[\int k(t)u(t - t_j)\,dt\right] \tag{12-31}$$

Now we can proceed as we did for the derivation of Eq. (12-17) and introduce the notion that the exact time at which a pulse occurs is randomly distributed with a uniform distribution function over the interval $0 \leq t \leq T$. If we suppose that there are precisely n pulses in this interval, the characteristic functional becomes

$$\Phi = \left(\frac{\gamma}{T}\right)^n \tag{12-32}$$

where

$$\gamma = \int W\left[\int k(t)u(t - s)\,dt\right] ds \tag{12-33}$$

If again we assume, as we did in the derivation of Eq. (12-18), that the distribution of the times of occurrences of the pulses fits the assumptions of the Poisson distribution, then we must multiply the expression of Eq. (12-32) by the factor $\bar{n}^n e^{-\bar{n}}/n!$, where, as before, $\bar{n} = \mu T$ is the expected number in the time interval T. We then sum on n to get

$$\begin{aligned}\Phi &= e^{-\mu(T-\gamma)}\\ &= \exp\left(-\mu\int\left\{1 - W\left[\int k(t)u(t - s)\,dt\right]\right\} ds\right)\end{aligned} \tag{12-34}$$

As a special example of this result, we assume that the pulse shape is extremely narrow. In fact, we assume that we can approximate the shape function by a delta function, that is, $u(t) = \delta(t)$. Then the characteristic functional is

$$\Phi = \exp\,(-\mu\textstyle\int\{1 - W[k(s)]\}\,ds) \tag{12-35}$$

Next, we assume that the distribution of scale heights is gaussian with zero mean and a root-mean-square value of σ; in other words, the ordinary normal distribution is given by

$$p(a)\,da = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-a^2/2\sigma^2}\,da \tag{12-36}$$

In that case, the characteristic function $W[\omega]$ becomes

$$W[\omega] = e^{-\sigma^2\omega^2/2} \tag{12-37}$$

and for Φ there results

$$\Phi[k(t)] = \exp\,[-\mu\textstyle\int(1 - e^{-(\sigma^2/2)[k(s)]^2})\,ds] \tag{12-38}$$

So we find again that a characteristic functional $\Phi[k(t)]$ can be derived to fit our assumed conditions. At any stage in this derivation, approximations that would reduce this to a quadratic form may be valid. For example, in the case just described a small value of the root-mean-square scale height σ corresponds to weak signals. If, at the same time, the expected number of signals arriving in a time interval is not small, then Eq. (12-38) can be approximated quite well by

$$\Phi = \exp\left\{-\frac{\mu\sigma}{2}\int\,[k(t)]^2\,dt\right\} \tag{12-39}$$

A distribution like this is called *white noise.*

12-4 GAUSSIAN NOISE

The type of distribution whose characteristic functional is gaussian comes up in many situations, and we shall discuss it here.

We have been working with probability distributions which are gaussian, i.e., exponentials of second order in the defining functions. Although we arrived at this gaussian functional by making a second-order approximation to the exponential term introduced by our assumption of a Poisson distribution of random pulses, it is worth remarking that a number of physical processes actually seem so distributed by their nature. In traditional probability theory the normal, or gaussian, distribution fits physical phenomena which are the result of the combination of a large number of independent events occurring randomly. This is the conclusion of the central-limit theorem of probability theory.[1] The same conclusion applies to distribution functionals and results in the fact that many important cases for study of physical phenomena have gaussian distributions. For further reference, we write here the most general form of a gaussian characteristic functional as

$$\Phi = e^{i\int k(t)F(t)\,dt}e^{-\frac{1}{2}\int\int k(t)k(t')A(t,t')\,dt\,dt'} \tag{12-40}$$

The first factor in this expression can be removed by a shift of the origin defining $f(t)$, as we discussed in deriving the distribution of fluctuations of voltage around a d-c level. Thus, we could define $f' = f - F(t)$. Next we note that, if the system we are describing behaves in a manner independent of the absolute value of time, then the kernel $A(t,t')$ must have the form $A(t - t')$.

In actual physical situations this particular function A may be defined by mechanisms in some sort of experimental situation or by approximating a particular piece of reality in such a way that it behaves nearly like the distribution function we are studying. We have an example of such an approximation in the derivations given above on the noise spectrum. For it, $A(t,t') = \mu\lambda(t - t')$. In either case theorems of the behavior of the system which result from the use of this function will be the same so long as the characteristic functional Φ can be suitably approximated by the quadratic or gaussian form of Eq. (12-40).

Of course, by now we know how to deal with gaussian functionals, since we have spent quite a bit of time in the preceding chapters manipulating them in one way or another. In this particular case the appearance of the factor i is different from that in typical quantum-mechanical cases. This means that functions which were real in Sec. 7-4, for example, are imaginary here. However, this does not require any review of the mathematical aspects of the subject; it simply is an awareness of and preparation for certain differences in detail in the results.

The probability distribution which corresponds to the characteristic functional of Eq. (12-40) is

$$P[f(t)] = \exp\{-\tfrac{1}{2}\int\int[f(t) - F(t)][f(t') - F(t')]B(t,t')\,dt\,dt'\} \tag{12-41}$$

---

[1] *Ibid.*, pp. 213ff.

where now the function $B(t,t')$ is a kernel reciprocal to $A(t,t')$. That is, the functions A and B are related by

$$\int A(t,\tau)B(\tau,s)\,d\tau = \delta(t-s) \tag{12-42}$$

Problem 12-1 Prove this.

All the parameters of the distribution can be calculated from the characteristic functional by the methods introduced in Chap. 7.

We shall now study in more detail some of the physical characteristics of gaussian noise that is time-independent; i.e., we shall study distributions whose characteristic functional is

$$\Phi = e^{-\frac{1}{2}\int\int k(t)k(t')A(t-t')\,dt\,dt'} \tag{12-43}$$

This function $A(\tau)$ is called the correlation function. Equation (12-43) means that the probability of observing a particular noise function $f(t)$ is

$$P[f(t)] = e^{-\frac{1}{2}\int\int f(t)f(t')B(t-t')\,dt\,dt'} \tag{12-44}$$

The function B appearing in this last expression is the *inverse* of the correlation function A. That is, $\int B(t-s)A(s)\,ds = \delta(t)$, or, if

$$\mathcal{P}(\omega) = \int A(\tau)e^{i\omega\tau}\,d\tau \tag{12-45}$$

is the Fourier transform of $A(\tau)$, the Fourier transform of $B(\tau)$ is $1/\mathcal{P}(\omega)$.

We shall begin by calculating some of the properties of this distributional. We first show that the average value of the noise signal vanishes. That is because the average value of the noise function at a particular time $t = a$ is, as in Eq. (12-13),

$$\langle f(a)\rangle = -i\frac{\delta\Phi}{\delta k(a)} \tag{12-46}$$

In this expression, the functional derivative of Φ in Eq. (12-43) is given by (cf. Sec. 7-2)

$$\frac{\delta\Phi}{\delta k(a)} = \left[-\int k(t)A(t-a)\,dt\right]\Phi \tag{12-47}$$

and, if it is evaluated for the particular function $k(t) = 0$, then it becomes 0.

Next we calculate the average of the square of the noise function or, better, the expected value of the product of two noise functions at times a and b. This is called the correlation function of the noise. It is [by differentiating both sides of Eq. (12-12) twice]

$$\langle f(a)f(b)\rangle = -\frac{\delta^2\Phi}{\delta k(a)\,\delta k(b)} = A(b-a)\Phi - \left[\int k(t)A(t-a)\,dt\right]\left[\int k(t')A(t'-b)\,dt\right]\Phi \qquad (12\text{-}48)$$

and, if this is evaluated for the function $k = 0$, it is simply $A(b-a)$. That is why A is called the correlation function.

12-5 NOISE SPECTRUM

A most useful characteristic of the noise distribution is the power spectrum of the noise (cf. Prob. 6-26), which is defined as the mean value of the square of the Fourier transform of the noise function, that is, the mean square of

$$\phi(\omega) = \int f(t)e^{i\omega t}\,dt \qquad (12\text{-}49)$$

By using our previous results, we can evaluate this as

$$\begin{aligned}\langle|\phi(\omega)|^2\rangle &= \left\langle\int f(a)e^{i\omega a}\,da \int f(b)e^{-i\omega b}\,db\right\rangle\\ &= \int\int\langle f(a)f(b)\rangle e^{i\omega(a-b)}\,da\,db\\ &= \int\int A(b-a)e^{i\omega(a-b)}\,da\,db\\ &= \int\mathcal{P}(\omega)\,da \end{aligned}\qquad (12\text{-}50)$$

Here we have made use of the function $\mathcal{P}(\omega)$, the Fourier transform of the correlation function A [cf. Eq. (12-45)].

If we carried out the integration defined in the last step of Eq. (12-50), we would, of course, get an infinite result. Therefore, the mean-square value which we are attempting to work out can be defined only for some finite time interval. If we take a unit time interval, then we say that the mean power per second is

$$\text{Mean of } |\phi(\omega)|^2 \text{ per second} = \mathcal{P}(\omega) \qquad (12\text{-}51)$$

We can apply some of these general results to our special example of noise produced by a multitude of small pulses. The correlation function for our problem is $\mu\lambda(\tau)$ introduced in Eq. (12-22). That is,

$$A(\tau) = \mu\int g(t)g(t+\tau)\,d\tau \qquad (12\text{-}52)$$

This means that the power function, actually called the power spectrum since it is defined in terms of frequency, is

$$\mathcal{P}(\omega) = \mu\int g(t)g(t+\tau)e^{i\omega\tau}\,d\tau\,dt = \mu|\gamma(\omega)|^2 \qquad (12\text{-}53)$$

where $\gamma(\omega)$ is the Fourier transform of our pulse function $g(t)$. We can explain this simple result more directly for our problem as follows. If the pulses occur at times t_i, so $f(t) = \sum_i g(t - t_i)$, the Fourier transform of $f(t)$ is $\phi(\omega) = \sum_i \gamma(\omega)e^{i\omega t_i}$. Thus the square of $\phi(\omega)$ has the average

$$\langle|\phi(\omega)|^2\rangle = \Big|\sum_{i,j} |\gamma(\omega)|^2 e^{i\omega(t_i - t_j)}\Big| \tag{12-54}$$

But, since the times t_i are at random and independent of t_j for $j \neq i$, all the terms for $i \neq j$ average out, because the average of $e^{i\omega(t_i - t_j)}$ is zero. Only the terms with $i = j$ remain. Each is $|\gamma(\omega)|^2$, and they are μT in number; so the mean of $|\phi(\omega)|^2$ per second is $\mu|\gamma(\omega)|^2$.

In the special case that the characteristic function can be approximated by the white-noise characteristic of Eq. (12-25), the function $A(t - t') = \text{const}\ \delta(t - t')$. This means that $\mathcal{P}(\omega)$ is independent of ω and there is the same "power" per unit frequency range [mean $|\phi(\omega)|^2$ per second] at all frequencies.

The distributions we are describing can very conveniently be described by giving the probability distribution not for $f(t)$ but for its Fourier transform $\phi(\omega)$ directly and the characteristic functional not in terms of $k(t)$ but its Fourier transform $\mathrm{K}(\omega)$.

$$\mathrm{K}(\omega) = \int k(t)e^{-i\omega t}\, dt \tag{12-55}$$

Using these functions, the characteristic functional for the noise distribution corresponding to Eq. (12-43) is

$$\Phi = e^{-\frac{1}{2}\int\int|\mathrm{K}(\omega)|^2\mathcal{P}(\omega)\, d\omega/2\pi} \tag{12-56}$$

by direct substitution of the inverse of Eq. (12-55) into Eq. (12-43). The corresponding probability functional of Eq. (12-26) is

$$P = e^{-\frac{1}{2}\int\int|\phi(\omega)|^2[1/\mathcal{P}(\omega)]\, d\omega/2\pi} \tag{12-57}$$

We can deduce the result of Eq. (12-57) from Eq. (12-56) very directly in the following way. We note that

$$e^{i\int k(t)f(t)\, dt} = e^{i\int \mathrm{K}(\omega)\phi(\omega)\, d\omega/2\pi} \tag{12-58}$$

so that Eq. (12-14) implies

$$P = \int \Phi e^{-i\int \mathrm{K}(\omega)\phi(\omega)\, d\omega/2\pi}\, \mathfrak{D}\mathrm{K}(\omega) \tag{12-59}$$

If we now imagine the possible values of ω to be discrete and separated by an infinitesimal spacing Δ, the integrals in the exponent in Eqs. (12-56) and (12-57) can be replaced by Riemann sums, and our path integral becomes

$$P = \prod_{\omega} \int e^{-(\Delta/2)|\mathrm{K}(\omega)|^2 \mathcal{P}(\omega)} e^{i\Delta \mathrm{K}(\omega)\phi(\omega)} \, d\mathrm{K}(\omega) \tag{12-60}$$

The integral for each value of ω can be done separately (by completing the square), and we get

$$\prod_{\omega} e^{\frac{-(\Delta/2)|\phi(\omega)|^2}{\Phi(\omega)}} \tag{12-61}$$

Putting the product together gives Eq. (12-57). It is clear that what happens at one frequency is independent of what happens at another, and the signal strength $\phi(\omega)$ at frequency ω is distributed as a gaussian with a mean square proportional to $\mathcal{P}(\omega)$.

12-6 BROWNIAN MOTION

It is usually true that the path integral method does not really help to get the solution to problems that cannot be solved in some other manner. Nevertheless, someone who has followed us this far and who is now familiar with path integrals will find its mode of expression and logic very simple and direct when applied to probability problems.

For example, in the theory of brownian motion we might have a linear system—say, a damped harmonic oscillator being driven by a fluctuating force $f(t)$. Assume the mass of the oscillator equal to 1, and we must solve

$$\ddot{x} - \gamma\dot{x} + \omega_0^2 x = f(t) \tag{12-62}$$

where $x(t)$ is the coordinate of the oscillator. If the function $f(t)$ is not known but is given by a known probability distribution $P_f[f(t)]$, what is the probability distribution $P_x[x(t)]$ for the various responses $x(t)$? Equation (12-62) relates x to f; that is, for each $f(t)$ there is an $x(t)$. Hence the probability of given x's is the same as that for the corresponding f's, or

$$P_x[x(t)]\,\mathfrak{D}x(t) = P_f[f(t)]\,\mathfrak{D}f(t) \tag{12-63}$$

where x is related to f via Eq. (12-62). In general, we must be very careful of the relation of $\mathfrak{D}x(t)$ to $\mathfrak{D}f(t)$, there being an analogue of a "jacobian" between the "volume" elements. But if f and x are linearly related (as above), this jacobian is a constant; so if, as is usual with path integrals, we can trust ourselves to be able to normalize our answer in the end, we have

$$P_x[x(t)] = \text{const } P_f(\ddot{x} - \gamma\dot{x} + \omega_0^2 x) \tag{12-64}$$

which gives us a formal solution. If P_f is gaussian, then P_x is and the problem may be worked out in many ways, the most evident being by the method of Fourier series if ω_0^2 and γ are independent of time.

At any rate, many problems can be set up and solved or partly solved by using Eq. (12-64) as a starting point. We shall look at a specific example. A fast particle goes through matter in which it receives small, sharp alterations in velocity as a result of passage by nuclei. After going through a thickness T, what is the probability it will emerge a distance D from the origin (the extension of its original straight-line path) and will be moving with deflection angle θ as in Fig. 12-1?

We assume that the interactions cause no measurable loss in the longitudinal velocity of the particle and that the matter through which the particle is passing is homogeneous. Further, we assume that θ is always small and that the motion is the result of a large number of collisions each of which has a small effect. We assume that the expected number of collisions in the infinitesimal thickness dt is μ and that the deflection suffered in each collision is given by the angle Δ, which is governed by the probability distribution $p(\Delta)\,d\Delta$. We further assume that this probability distribution results in a mean-square value of Δ given by

$$\int_{-\infty}^{\infty} \Delta^2 p(\Delta)\,d\Delta = \sigma^2 \tag{12-65}$$

and we shall use the substitution $\mu\sigma^2 = R$.

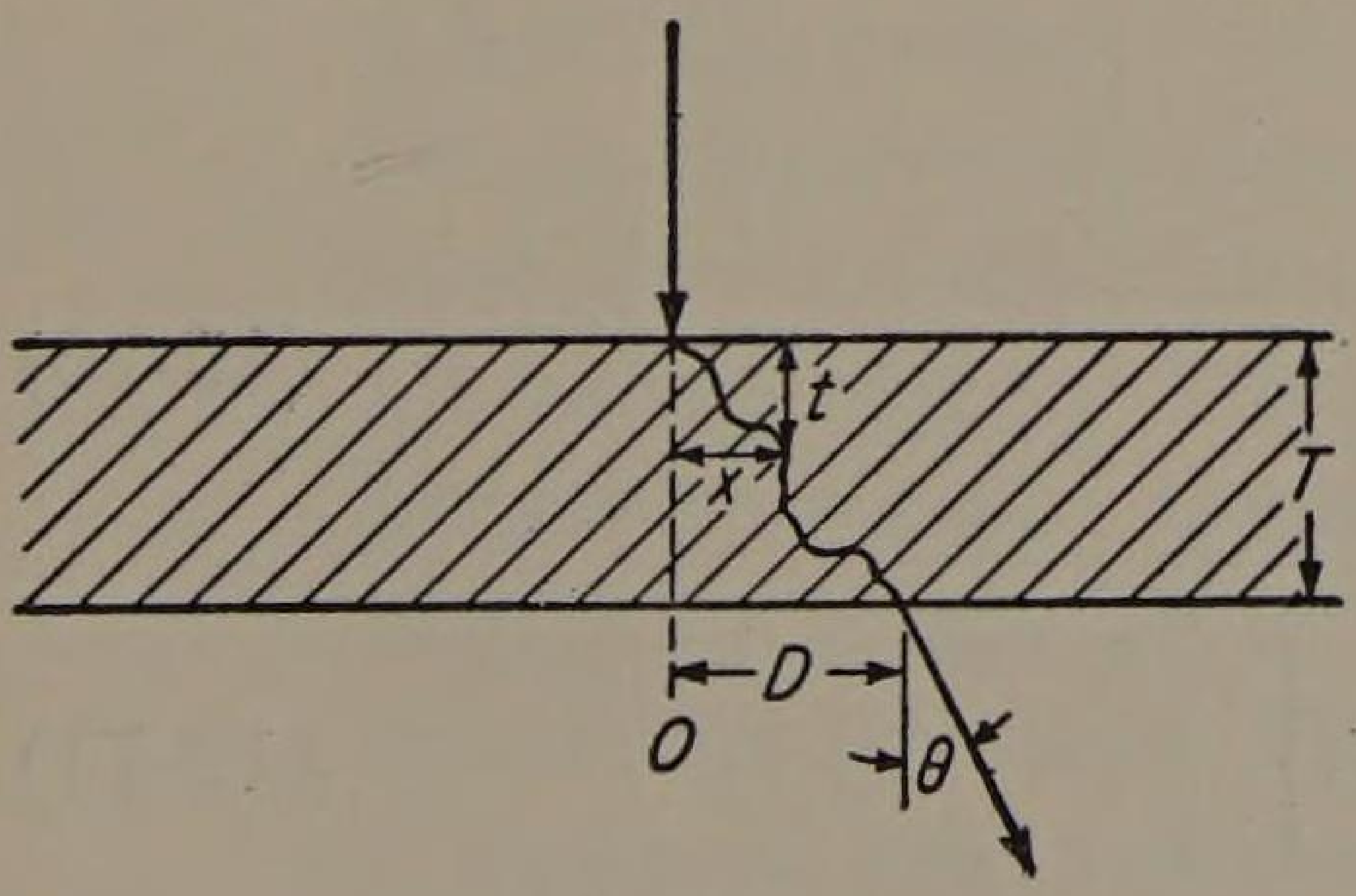


Fig. 12-1 A fast particle impinges perpendicularly on a slab of matter of thickness T. After traveling through a thickness t measured parallel to its original line of flight, it is deflected away from its original trajectory (as extended) by a distance x owing to a number of interactions with the nuclei in the material. Eventually, it emerges from the slab a distance D from the point O, at which it would have emerged with no deflection, and is traveling in a direction that makes an angle θ with its original direction.

We shall confine our attention to the motion as projected onto a two-dimensional plane containing the original path of the particle. Motion in a plane normal to this will follow similar rules, and the motion in either plane can be considered independently of the other. We shall use t to measure the depth of penetration into the slab, θ to represent the instantaneous direction of motion in the plane we are considering, and x to measure the position of the particle away from an extension of its original path of motion, as shown in Fig. 12-1. These parameters are related by $dx = \theta\, dt$, or $\dot{x} = \theta$.

We assume that the deflections of θ occur suddenly, so that $\dot{\theta} = f(t)$, where the functions f are a set of randomly spaced delta functions having random scale heights. This means that $\ddot{x} = f(t)$ and $P_f[f(t)]$ has the characteristic functional

$$\Phi = e^{-\mu\int\{1-W[k(s)]\}\,ds} \tag{12-66}$$

where

$$W[\omega] = \int p(\Delta) e^{i\omega\Delta}\, d\Delta \tag{12-67}$$

We note that the mean value of Δ is assumed to be 0, and these deflections themselves are assumed small. Now if we expand $G(w)$ as

$$W[\omega] = \int p(\Delta)\left(1 + i\omega\Delta - \frac{\omega^2}{2}\Delta^2 + \cdots\right) d\Delta \tag{12-68}$$

and use terms only through second order in Δ to get $W[\omega] = 1 - \omega^2\sigma^2/2$, then

$$\Phi = e^{-\frac{1}{2}R\int[k(s)]^2\,ds} \tag{12-69}$$

This in turn implies that

$$P_f[f(t)] = e^{-(1/2R)\int[f(t)]^2\,dt} \tag{12-70}$$

Hence

$$P_x[x(t)] = \text{const}\ \exp\left\{-\frac{1}{2R}\int_0^T [\ddot{x}(t)]^2\, dt\right\} \tag{12-71}$$

We wish to evaluate the probability distribution $P(D,\theta)$, which gives the probability that the particle will exit with the displacement D and angle of motion θ under the assumptions that it enters with the initial conditions $x(0) = 0$ and $\dot{x}(0) = 0$. We are concerned not with the exact path that the particle takes in the material, but only that the particle exits with the conditions $x(T) = D$ and $\dot{x}(T) = \theta$. Thus, we express this probability distribution by an integral over all paths as

$$P(D,\theta) = \int \exp\left(-\frac{1}{2R}\int_0^T \ddot{x}^2\,dt\right)\mathcal{D}x(t) \tag{12-72}$$

where the paths included in the integral satisfy the assumed end-point conditions. This integral can be carried out by the methods of Sec. 3-5. The integral is a gaussian and becomes an extremum for the path

$$\ddot{\ddot{x}}(t) = 0 \tag{12-73}$$

The solution of this equation, which satisfies our assumed boundary conditions, is

$$x(t) = (3D - \theta T)\left(\frac{t}{T}\right)^2 + (\theta T - 2D)\left(\frac{t}{T}\right)^3 \tag{12-74}$$

By using this path in the integrand of the exponential function in Eq. (12-72), we find

$$\frac{1}{2R}\int_0^T \ddot{x}^2\,dt = \frac{6}{RT^3}\left(D - \frac{\theta T}{2}\right)^2 + \frac{\theta^2}{2RT} \tag{12-75}$$

which means that our required probability distribution is

$$P(D,\theta) = \text{const}\exp\left[-\frac{6}{RT^3}\left(D - \frac{\theta T}{2}\right)^2 - \frac{\theta^2}{2RT}\right] \tag{12-76}$$

In some practical cases we may really be concerned not with the exact linear spacing of the particle away from our assumed origin point but, rather, with the deflection angle at which it leaves the slab. Given the overall distribution function of Eq. (12-76), it is simple to evaluate the distribution function in angle alone by integrating over all values of D. The result is $e^{-(\theta^2/2RT)}$. This is an expected result, because we have already assumed that the mean-square value of the deflection angle which would be acquired in a unit thickness is R, so this value in a total thickness T should be RT.

Suppose next we look only at particles which emerge traveling in a specific angle θ and consider the distribution function of the emerging positions D of those particles. We find that the probability distribution has a maximum at $D = \theta T/2$. This would be the position we would expect if the final deflection angle θ were acquired in a smooth manner as a linear function of thickness starting from 0 and building up to its final value. In that case its average value during the passage through the slab would be $\theta/2$.

Problem 12-2 Show that the constant required to normalize the probability function $P(D,\theta)\,dD\,d\theta$ is

$$\text{const} = \sqrt{\frac{6}{\pi R T^3}}\sqrt{\frac{1}{2\pi R T}} \tag{12-77}$$

12-7 QUANTUM MECHANICS

In this and the following sections we should like to see how to formulate statistical problems in quantum mechanics. In quantum mechanics there are probabilities involved in an intrinsic way, because even a known state implies probabilities to be found in other states. But in addition there may be extrinsic uncertainties. The state, for example, may not be known—we may know only that the state is such and such with a certain probability. This situation is analogous to the classical-mechanics situation in which the initial conditions are not known and only a probability distribution for such conditions is available. We have already dealt with such a situation in statistical mechanics (cf. Chap. 10), but that is a very special case in which the state of energy E has the probability $e^{-E/kt}$. Here we shall be more general.

Again, under a given external potential, say $V(t)$, the behavior of a quantum-mechanical system can be worked out, but what can we say if that potential is uncertain and has a probability distribution $P[V(t)]\,\mathfrak{D}V(t)$? Need we actually solve the problem for each $V(t)$ and then average, or is there some way to formulate the problem *after* the average of $V(t)$ is taken? (We hope so, because it often occurs that the solution of a statistical problem after an average is taken is, in fact, much easier than finding the general solution of the original problem for a wide range of conditions.) We shall find such a formulation in this section. Then we shall go on to discuss situations in which a quantum-mechanical system is disturbed not just by a classical system but by another quantum-mechanical system about which there are statistical uncertainties.

Our main purpose in this chapter is to show how these and other questions may be formulated. We shall not deal in detail with solving the special problems mentioned; they are brought up only to help us understand the more general formulations we shall arrive at.

We wish first to discuss the analogue of brownian motion for a quantum-mechanical system. That is, we shall suppose that a quantum-mechanical system whose unperturbed action is $S(q)$ is under the influence of an external potential $V(t)$ such that its action is[1]

$$S_V(q) = S(q) + \int q(t)V(t)\,dt \tag{12-78}$$

Suppose we ask: What is the probability that, starting at some early time t_i with the coordinate $q(t_i) = q_i$, we shall arrive at a final time t_f at position q_f? It is the square of an amplitude: $|K(q_f,t_f;q_i,t_i)|^2$. Or again, if we specify that initially a system be in a state of wave function $\phi(q)$ and finally in wave function $\chi(q)$, the probability of transition from ϕ to χ is

$$\begin{aligned} P[\chi(q);\phi(q)] &= \left|\int\int \chi^*(q_f)K(q_f,t_f;q_i,t_i)\phi(q_i)\,dq_f\,dq_i\right|^2 \\ &= \int\int\int\int \chi^*(q_f)\chi(q'_f)K(q_f,t_f,q_i,t_i)K^*(q'_f,t_f;q'_i,t_i)\phi(q_i)\phi^*(q'_i) \\ &\qquad dq_i\,dq'_i\,dq_f\,dq'_f \end{aligned} \tag{12-79}$$

It is evident that all such problems can be solved if we can evaluate

$$K(q_f,t_f;q_i,t_i)K^*(q'_f,t_f,q'_i,t_i) \tag{12-80}$$

The first factor involves the path integral $\int e^{iS[q(t)]}\,\mathfrak{D}q(t)$, whereas the second complex conjugate[2] one is $\int e^{-iS[q(t)]}\,\mathfrak{D}q(t)$. Each integral is over paths with appropriate end points. In writing the product of Eq. (12-80), we shall call the path variable in the second integral $q'(t)$ and we can then express Eq. (12-80) as the *double path integral*

$$\int\int e^{iS[q(t)]-iS[q'(t)]}\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t) \tag{12-81}$$

The summing of such integrals over various end points gives the required probability.

If the potential V is acting, we should replace S in Eq. (12-81) by S_V, and the expression becomes

$$\int\int e^{i\{S[q(t)]-S[q'(t)]+\int q(t)V(t)\,dt-\int q'(t)V(t)\,dt\}}\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t) \tag{12-82}$$

But now suppose the potential is known only in a probabilistic sense; i.e., we know that there is the probability $P_V[V(t)]\,\mathfrak{D}V(t)$ that the potential is $V(t)$. Then the probability to go from ϕ to χ is given by

---

[1] We shall do everything as though there were only one coordinate q. One can immediately generalize to several coordinates q_i (so V is a set of potentials V_i) and to cases in which the coefficient of $V(t)$ in the action is not simply q but some more complex operator.

[2] We suppose that $S[q(t)]$ is real and that our units are so defined that $\hbar = 1$, as in Chap. 11.

Eq. (12-79) calculated for each $V(t)$ and then averaged over all $V(t)$ each with the weight $P_V[V(t)]\,\mathfrak{D}V(t)$. This is then

$$\text{Prob}\,(\phi \rightarrow \chi) = \int\!\!\int\!\!\int\!\!\int \chi^*(q_f)\chi(q_f')J(q_f,q_f';q_i,q_i')\phi(q_i)\phi^*(q_i')\,dq_i\,dq_i'\,dq_f\,dq_f' \tag{12-83}$$

where J is the average of Eq. (12-82) over all $V(t)$ with weight $P_V[V(t)]\,\mathfrak{D}V(t)$; thus

$$J = \int\!\!\int\!\!\int e^{i\{S[q(t)]-S[q'(t)]\}}e^{i\int[q(t)-q'(t)]V(t)\,dt}P_V[V(t)]\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t)\,\mathfrak{D}V(t) \tag{12-84}$$

with the integrals taken between appropriate end points $q(t_i) = q_i$, $q'(t_i) = q_i'$, $q(t_f) = q_f$, $q'(t_f) = q_f'$. Actually, this choosing of end points and then integrating over various values with wave-function distributions depending on the problem [as in Eq. (12-83)] is simply a sum of J's for different end conditions, and we shall hereafter simply forget this and speak as though with J we already have our probability —it being left to the reader to remember that a bit more has yet to be done. This is so that we can concentrate on the main feature, the evaluation of the double path integrals needed to calculate J.

In this form we can do the integral over $V(t)$ explicitly and see that, to find the probabilities after averaging, we must evaluate a double path integral

$$J = \int\!\!\int e^{i\{S[q(t)]-S[q'(t)]\}}\Phi[q(t) - q'(t)]\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t) \tag{12-85}$$

where $\Phi[k(t)]$ is the generating functional belonging to the probability distribution P_V, so

$$\Phi[k(t)] = \int e^{i\int k(t)V(t)\,dt}P_V[V(t)]\,\mathfrak{D}V(t) \tag{12-86}$$

Equation (12-85) then answers our challenge to express the answer in a form valid after the averaging. It involves evaluation of the double path integral. How to evaluate it is, of course, another question, but the methods discussed in this book may be useful. In these sections we are discussing only how various problems may be formulated.

As an example of the application of Eq. (12-85), suppose $V(t)$ is gaussian noise with zero mean and characteristic function $A(t,t')$ as in Eq. (12-46). We must evaluate

$$J = \int\!\!\int \exp i\{S[q(t)] - S[q'(t)]\} \exp\{-\tfrac{1}{2}\int\!\!\int[q(t) - q'(t)] \times [q(t') - q'(t')]A(t,t')\,dt\,dt'\}\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t) \tag{12-87}$$

Because in the new factor at least the q and q' appear only quadratically, some of the methods previously discussed for quadratic forms may be useful. Of course, if $S(q)$ is itself quadratic, corresponding to a harmonic oscillator, the path integrals can be evaluated exactly by using the methods of Sec. 3-5.

12-8 INFLUENCE FUNCTIONALS

Now we wish to discuss the behavior of a quantum-mechanical system whose general coordinate we shall take to be q in interaction with another system whose coordinate we call Q.† We shall suppose that all measurements which are to be made are on system q only, and no direct measurements of the system Q will be made. For example, we may be interested in how an atom makes transitions because it is in the electromagnetic field and can radiate. We contemplate studying only the atom and will not directly measure the light coming from it; then q are the atomic coordinates and Q the coordinates of the field. If we study it the other way—that is, if we only observe the light from the atom, emitted, absorbed, or scattered, but never ask for any quantity directly involving the atom's variables—then we may use our present analysis with Q being the atomic coordinates and q those of the electromagnetic field. If, for example, the theory of the index of refraction is wanted, then q are again the field variables and Q the variables of the piece of matter through which the light goes. For one further example, suppose the behavior of an electron in a crystal (or an ion in a liquid) is to be studied: the measurements to be analyzed involve directly only the position of the charge, not the material of the crystal. For example, we might wish the current (electron velocity) generated in some circumstance, but we are not contemplating correlations with the number of phonons produced. Then q can be the coordinates of the electron and Q all the other coordinates of the matter of the crystal.

Let $S[q(t)]$ be the action of system q, $S_0[Q(t)]$ that of the environmental system alone, and $S_i[q(t),Q(t)]$ that of the interaction between the environmental system Q and the system of interest q. The action of the combined system is $S[q(t)] + S_0[Q(t)] + S_i[q(t),Q(t)]$, and the probability of any event involving the combined system can be evaluated from the double path integral, an obvious generalization of Eq. (12-81), and now written as

---

† Q stands for any number of coordinates—this other system may be, and generally is, very complex. We shall just carry one Q variable, but nothing essential will be lost.

$$J = \iint \exp\,(i\{S[q(t)] - S[q'(t)] + S_{int}[q(t),Q(t)] - S_{int}[q'(t),Q'(t)] + S_0[Q(t)] - S_0[Q'(t)]\})\,\mathfrak{D}q(t)\,\mathfrak{D}Q(t)\,\mathfrak{D}q'(t)\,\mathfrak{D}Q'(t) \tag{12-88}$$

But, if we need no measurements on $Q(t)$ and if only the dependence on $q(t)$ need ever be studied, then we can write our answer in the form

$$J = \iint e^{i\{S[q(t)]-S[q'(t)]\}}F[q(t),q'(t)]\,\mathfrak{D}q(t)\,\mathfrak{D}q'(t) \tag{12-89}$$

where we shall call the functional $F[q(t),q'(t)]$ the *influence functional.* It is a functional of the two functions $q(t)$ and $q'(t)$, and for this particular problem it is given by

$$F[q(t),q'(t)] = \sum_f \iint \exp\,(i\{S_0[Q(t)] - S_0[Q'(t)] + S_{int}[q(t),Q(t)] - S_{int}[q'(t),Q'(t)]\})\,\mathfrak{D}Q(t)\,\mathfrak{D}Q'(t) \tag{12-90}$$

The sum on f means we are to sum over all possible final states of Q. This is because no measurement on Q is to be taken, and all final states of the environment are possible. Therefore we must add together the probabilities [i.e., the J functions of Eq. (12-88)] of all. In coordinate representation, for example, $\sum_f$ just means that at some final time t_f after we are no longer interested in the interaction we must take $Q(t_f) = Q'(t_f) = Q_f$ and integrate over all Q_f.

To summarize, the behavior of a system in any environment can be discussed in terms of a double path integral like Eq. (12-89), where F is a property of the environment—its "influence" on the system. It summarizes all of the environment that is relevant to $q(t)$. Two different possible surrounding conditions, say, A and B, might physically be very differently constructed; nevertheless, if they happen to lead to the same functional F, they are indistinguishable as far as the behavior of the q system is concerned.

This F is somewhat analogous to the use of "external force" in separating the behavior of interacting systems classically. We can analyze the motion of q alone provided we suppose we know what force is produced (as a function of time) by the environment. These newtonian equations of motion of q alone are the rough analogue of Eq. (12-89), whereas Eq. (12-90) corresponds to the calculation of the force produced by a given environment. Two different environments which produce the same force on q are equivalent. Actually, the analogy is only rough; for F contains the entire effect of the environment including the change in behavior of the environment resulting from reaction with q. In the classical analogue, F would correspond to knowing not only what the force is as a function of time, but also what it would be for every possible motion $q(t)$ of the object. The force for a given environmental system depends in general on the motion of $q(t)$, of course, since the environmental system is affected by interaction with the system of interest q.

We are therefore led to study the properties of influence functionals. We shall be content to list a few such rules and give some suggestions on how they are arrived at.

Rule I

$$F[q(t),q'(t)]^* = F[q'(t),q(t)] \tag{12-91}$$

where the asterisk means complex conjugate.

Rule II. If the argument functions $q(t)$ and $q'(t)$ are set equal for t exceeding any value a, then F does not depend upon the actual values of $q(t)$ for $t > a$.

Rule III. If F_i is the influence functional for a particular environment i and we do not know what the environment actually is but know only that the probability of its being i is ω_i, then the effective influence functional (for calculating all probabilities) is

$$F = \sum_i \omega_i F_i \tag{12-92}$$

Rule IV. If the system q is simultaneously in interaction with two external systems A and B and if A and B do not interact directly with each other and there is no correlation between them in the initial conditions, then

$$F = F_A \cdot F_B \tag{12-93}$$

where F_A is the influence functional if A alone were interacting and F_B is that if B alone were interacting.

Rule V. If the functional F can be adequately approximated by the form

$$F = \exp\{i\int[q(t) - q'(t)]V(t)\,dt\} \tag{12-94}$$

the system q is acting as though under a classical potential $V(t)$ with action of interaction $\int q(t)V(t)\,dt$. If it is of the form $F(q,q') = \Phi[q(t) - q'(t)]$, where $\Phi[k(t)]$ is any functional, the environment is equivalent to a classical but uncertain potential $V(t)$. [Φ is the characteristic functional for the distribution of $V(t)$.]

That rule I is true is evident directly from Eq. (12-90). This expression also explains rule II, but in a much more subtle way. Note that for any system with any definite action $S_a(Q)$ and any given initial state

$$\sum_f \iint e^{i\{S_a[Q(t)]-S_a[Q'(t)]\}}\,\mathfrak{D}Q(t)\,\mathfrak{D}Q'(t) = 1 \tag{12-95}$$

This is because the integrals and the sum over final states $\sum_f$ are equivalent to

$$\int K(Q_f,t_f;Q_i,t_i)K^*(Q_f,t_f;Q'_i,t_i)\,\mathfrak{D}Q_f = \delta(Q_i - Q'_i) \tag{12-96}$$

by Eq. (4-37). Thus, if the initial wave function were $\phi(Q_i)$, we would multiply by $\phi(Q_i)\phi^*(Q_i)$ as we did in Eq. (12-79) and integrate to get

$$\begin{aligned}\int \phi(Q_i)\phi^*(Q'_i)\delta(Q_i - Q'_i)\,dQ_i\,dQ'_i &= \int |\phi(Q)|^2\,dQ\\ &= 1\end{aligned} \tag{12-97}$$

Now notice that, if we put $q'(t) = q(t)$ for a given fixed $q(t)$ for all time in Eq. (12-90), we have an expression just like Eq. (12-95) where the effective (and definite) action is

$$S_a[Q(t)] = S_0[Q(t)] + S_{int}[q(t),Q(t)]$$

with

$$S_a[Q'(t)] = S_0[Q'(t)] + S_{int}[q(t),Q'(t)]$$

as required, as long as $q'(t) = q(t)$. Hence $F[q(t),q(t)] = 1$.

The same argument limited to the time range $a \leq t \leq t_f$, using a relation like Eq. (12-96) but with t_i, Q_i replaced by a, Q_a, shows that, if $q(t) = q'(t)$ for $t > a$, the dependence of F on $q(t)$ for $t > a$ drops away, because the right side of Eq. (12-96) does not depend on $q(t)$ for $t > a$.

Rule III is an evident result of the fact that probabilities are determined by adding the value of J over various circumstances.

Rule IV is evident from Eq. (12-90) when it is realized that the conditions of the rule imply that the action that goes into Eq. (12-90) is $S_{0A}[Q_A(t)] + S_{int\ A}[q(t),Q_A(t)] + S_{0B}[Q_B(t)] + S_{int\ B}[q(t),Q_B(t)]$ and that the exponential of the sum becomes a product, as does the integral F, if the initial state is itself a product of wave functions.

Rule V is merely a statement of our results shown in Eqs. (12-82) and (12-85).

These are some of the general properties of influence functionals. Calculations with them involve the various methods for doing path integrals applied to Eq. (12-89). We shall conclude this section by discussing certain important influence functionals.

Just as gaussian probability distributions and gaussian noise distributions are simple and important, so influence functionals which depend on $q(t)$, $q'(t)$ as an exponential of a quadratic form—which we shall call gaussian influence functionals—are particularly important.

First, if the environment is a set of harmonic oscillators in their ground state (or at a given temperature) coupled linearly to the system of interest q, evaluation of Eq. (12-90) shows that F is gaussian. But gaussian influence functionals, like gaussian probabilities, occur in good approximation in a much wider class of situations, namely, where the effect is the result of a very large number of influences, each of which by itself has little effect. For example, consider an atom in weak interaction with each of the large number of atoms of an environmental gas. The influence of one atom A is very small, so its influence functional F_A differs only slightly from 1. However, in view of rule IV, the complete F is the product of many such factors, which becomes (nearly) the exponential of the sum of a small contribution from each. This contribution expanded to first and second order in the interaction with each atom leads to influence functionals of the gaussian type.

As an application of this conclusion, a piece of metal placed in a cavity resonator affects the resonator in a simple linear way summarizable by one impedance function, even though the multitude of electrons in the metal behave in such a complex manner. The influence functional of the metal (Q) on the cavity oscillator (q) is nearly a gaussian, and to this extent the metal is equivalent to some set of harmonic oscillators which would produce the same influence functional.

The most general exponential functional involving $q(t)$ and $q'(t)$ in linear form is

$$F[q(t),q'(t)] = \exp\left[i\int q(t)V(t)\,dt - i\int q'(t)U(t)\,dt\right] \tag{12-98}$$

for arbitrary and complex $V(t)$ and $U(t)$. If this is to be an influence functional, however, it must satisfy the conditions of our five rules. Rule I requires $U(t) = V^*(t)$, and rule II implies $U(t) = V(t)$; hence U and V are equal and V is real. Thus the most general linear functional is that equivalent to the action of a classical external potential in accordance with rule V.

We need not discuss this simple case further; for it is completely analyzable just by adding $q(t)V(t)$ to the hamiltonian of the unperturbed problem. If the exponent has both quadratic and linear terms, the linear term can be factored out, so via rule IV we can say it is a classical potential plus the effect of a purely quadratic functional.

The most general exponential functional which involves its arguments purely quadratically is of the form

$$F[q(t),q'(t)] = \exp\Big\{-\iint^{t} [\alpha(t,t')q(t)q(t') + \beta(t,t')q'(t)q'(t') + \gamma(t,t')q(t)q'(t') + \delta(t,t')q'(t)q(t')]\,dt'\,dt\Big\} \tag{12-99}$$

for arbitrary and complex[1] α, β, γ, and δ. The integrals on t', t are over the entire interesting range of time, but we always take $t > t'$. This is no loss of generality, of course, but it is convenient for later analysis. For this to be a satisfactory influence functional, we must have from Rule I

$$\beta(t,t') = \alpha^*(t,t') \tag{12-100}$$

and

$$\gamma(t,t') = \delta^*(t,t') \tag{12-101}$$

Rule II gives us a great deal of information, for putting $q(t) = q'(t)$ for $t > a$ and, assuming $t > a$, $t' < a$, the expression [which is part of the integral in Eq. (12-99)]

$$\int_a \int^a [\alpha(t,t')q(t)q(t') + \beta(t,t')q(t)q'(t') + \gamma(t,t')q(t)q'(t') + \delta(t,t')q(t)q(t')]\,dt\,dt' \tag{12-102}$$

must be independent of $q(t)$ for $t > a$ and arbitrary $q(t')$ and of $q'(t')$ for $t' < a$. This requires that

$$\begin{aligned} \delta(t,t') &= -\alpha(t,t') \\ \gamma(t,t') &= -\beta(t,t') \end{aligned} \tag{12-103}$$

as long as $t > a$, $t' < a$. But since a is arbitrary, Eqs. (12-103) must hold for all t, t' (under the continuing restriction $t > t'$).

Therefore, the most general gaussian influence functional depends on only one complex function $\alpha(t,t')$ and is of the form

$$\exp\Big\{-\iint^{t} [q(t) - q'(t)][q(t')\alpha(t,t') - q'(t')\alpha^*(t,t')]\,dt'\,dt\Big\} \tag{12-104}$$

In the case that $\alpha(t,t')$ is real, say, $A(t,t')$ our functional is equivalent to the exponential of Eq. (12-87), and we have the equivalent of a noisy classical perturbation. In general, in quantum-mechanical systems α is complex. A special case of importance is that

---

[1] These functions are defined only for $t > t'$.

$$\alpha(t,t') = \alpha(t - t')$$

depends only on the time difference $t - t'$. We are then dealing with an environmental system which has average properties independent of absolute time.

To help understand some of the properties of Eq. (12-104), we shall ask for the probability that the q system makes a transition from energy level n to some other orthogonal level m during an interval of time T in the case that α is very small and we can use perturbation theory. If we expand F in Eq. (12-104), the leading term, 1, gives nothing because the states are orthogonal. The next term linear in α has four pieces. One is $-\iint^t \alpha(t,t')q(t)q(t')\,dt'\,dt$. When this is substituted into Eq. (12-89) in place of F and this evaluated as in Eq. (12-83) with $\phi = \phi_n$ and $\chi = \phi_m$, the integral on $\mathfrak{D}q(t)$ and $\mathfrak{D}q'(t)$ is seen to be the product of two factors. One, the integral on q, involves

$$\int e^{iS[q]}\left[-\iint^t \alpha(t,t')q(t)q(t')\,dt'\,dt\right]\mathfrak{D}q(t)$$

which is the transition element (cf. Chap. 4)

$$\begin{aligned} {}_m\!\left\langle -\iint^t \alpha(t,t')q(t)q(t')\,dt'\,dt\right\rangle_n \\ = -\iint^t {}_m\langle q(t)q(t')\rangle_n\,\alpha(t,t')\,dt'\,dt \qquad (12\text{-}105)\end{aligned}$$

The integral on $\mathfrak{D}q'$ is just $\int e^{iS[q']}\,\mathfrak{D}q'$ and is the complex conjugate of the transition element ${}_m\langle 1\rangle_n$. Analyzing the other terms in a similar way, the total transition probability is

$$\begin{aligned} P(n \to m) = \iint^t [&-\alpha(t,t')\;{}_m\langle q(t)q(t')\rangle_n\;{}_m\langle 1\rangle_n^* \\ &- \alpha^*(t,t')\;{}_m\langle 1\rangle_n\;{}_m\langle q(t)q(t')\rangle_n^* + \alpha^*(t,t')\;{}_m\langle q(t)\rangle_n\;{}_m\langle q(t')\rangle_n^* \\ &+ \alpha(t,t')\;{}_m\langle q(t)\rangle_n^*\;{}_m\langle q(t')\rangle_n]\,dt'\,dt \qquad (12\text{-}106)\end{aligned}$$

If m and n are orthogonal, ${}_m\langle 1\rangle_n = 0$. If $S[q]$ comes from a constant hamiltonian with energy levels E_k for states k, then

$$ {}_m\langle q(t)\rangle_n = q_{mn}e^{-i(E_m-E_n)t} \qquad (12\text{-}107)$$

Only the last two terms of Eq. (12-106) survive, and they are complex conjugates of each other, so that

$$P(n \to m) = 2\,\text{R.P.}\iint^t \alpha(t,t')e^{-i(E_m-E_n)(t-t')}\,dt'\,dt \qquad (12\text{-}108)$$

Problem 12-3 For $m = n$, verify $P(m \rightarrow m) = 1 - \sum_n P(m \rightarrow n)$ as required by conservation of probability.

In the case of a time-steady environment $\alpha(t,t') = \alpha(t - t')$. Suppose we define the Fourier transform

$$a(\nu) = \int_0^\infty \alpha(\tau)e^{-\nu\tau}\, d\tau \tag{12-109}$$

[$\alpha(t)$ is not defined for $t < 0$.] Then since $P(n \rightarrow m)$ in Eq. (12-108) is proportional to the time interval over which the integrals extend, we can define a rate of transition per second and find the probability of transition

$$P(n \rightarrow m) \text{ per second} = 2a_R(E_m - E_n)|q_{nm}|^2 \tag{12-110}$$

where we break $a(\nu)$ into real and imaginary parts

$$a(\nu) = a_R(\nu) + ia_I(\nu) \tag{12-111}$$

We may note that, for a disturbance by a classical potential under gaussian noise, $\alpha(\tau)$ is real [cf. Eq. (12-87)] and the real part of $a(\nu)$ is the power-spectrum function of the noise as defined in Eq. (12-32). So, for such *classical noise* systems

$$a_R(\nu) = a_R(-\nu) \tag{12-112}$$

and in first-order perturbation

$$\text{Rate of transition } n \rightarrow m = \text{rate of transition } m \rightarrow n \tag{12-113}$$

and both rates are proportional to the power $P(\omega)$ at the frequency of the transition. Thus classical potentials have equal probability of causing transitions up and down.

Another interesting example is when the environment cannot supply energy with any reasonable probability. For example, it may be initially in the ground state or at zero temperature. We shall call such an environment "cold." For such a situation transitions of the system q going up in energy ($E_m > E_n$) are unlikely. Hence for such *cold-environment* systems

$$a_R(\nu) = 0 \qquad \text{for } \nu > 0 \tag{12-114}$$

and for first-order perturbations

$$\text{Rate of transition } n \rightarrow m = 0 \qquad \text{if } E_m > E_n \tag{12-115}$$

Since any $a(\nu)$ can be written as the sum of one of the type shown in Eq. (12-112) plus one of the type shown in Eq. (12-114), it is readily apparent that any time-independent gaussian functional is equivalent to a system in some cold environment acted on by a fluctuating classical potential described by a gaussian expression. This conclusion follows from the fact that the product of any two gaussian functions is also a gaussian and from rule IV. If the interaction of one environment on the system is represented by $A_1(t,t')$ in the manner of Eq. (12-87) and the interaction of the other environment as $A_2(t,t')$, then the single interaction term in the single resulting gaussian functional is $A_1 + A_2$.

12-9 INFLUENCE FUNCTIONAL FROM A HARMONIC OSCILLATOR

We shall next give an example of how F can be worked out from Eq. (12-90) for an environment consisting of a harmonic oscillator with coordinates Q, in the ground state, coupled to q linearly with an interaction $S_{int}(q,Q) = C\int q(t)Q(t)\,dt$. We suppose the oscillator of unit mass and frequency ω, so that

$$S_0(Q) = \tfrac{1}{2}\int[\dot{Q}(t)^2 + \omega^2 Q(t)^2]\,dt \tag{12-116}$$

Then

$$\begin{aligned}F[q(t),q'(t)] = \sum_m \iint \exp\left\{i\int [\tfrac{1}{2}\dot{Q}(t)^2 + \tfrac{1}{2}\omega^2 Q(t)^2\right.\\ \left. + Cq(t)Q(t)]\,dt\right\} \exp\left\{-i\int [\tfrac{1}{2}\dot{Q}'(t)^2 + \tfrac{1}{2}\omega^2 Q'(t)^2\right.\\ \left. + Cq'(t)Q'(t)]\,dt\right\} \mathcal{D}Q(t)\,\mathcal{D}Q'(t)\end{aligned} \tag{12-117}$$

where m is the final state and the initial state is the ground state. The integral on Q is clearly gaussian, and in fact we have already done it; for it is exactly the transition amplitude G_{m0} worked out in Sec. 8-9 for a forced harmonic oscillator. The forcing function there called $\gamma(t)$ is here $Cq(t)$.† It is therefore given by Eq. (8-145) with $n = 0$ or

$$G_{m0} = (m!)^{-1/2}(i\beta^*)^m G_{00} \tag{12-118}$$

---

† The reader may prefer to observe that Eq. (12-117) is

$$F[q(t),q'(t)] = \int dQ_f\, K(Q_f,t_f;Q_i,t_i)K'^*(Q_f,t_f;Q_i',t_i)\,\phi_0(Q_i)\phi_0^*(Q_i')\,dQ_i\,dQ_i'$$

where K is the kernel of Eq. (3-66) for a forced harmonic oscillator with $f(t) = Cq(t)$ and K' is that with $f(t) = Cq'(t)$. $\phi_0(Q)$ is the wave function of the oscillator in the ground state. All variables Q_i, Q_i', and Q_f then appear in a simple gaussian way and may be directly integrated. We shall then find it as easy to do the finite-with G_{00} given in Eq. (8-138) and β^* in Eq. (8-143) replacing γ by $Cq(t)$. Likewise, the integral over Q' is the complex conjugate of a similar expression but with $\gamma(t)$ replaced by $Cq'(t)$ this time. We distinguish values with this substitution with a prime. Then the sum on final states in Eq. (12-117) gives us

$$F(q,q') = \sum_m G_{m0}G'^*_{m0} = \sum_n (m!)^{-\frac{1}{2}}(i\beta^*)^m G_{00}(m!)^{-\frac{1}{2}}(-i\beta')^m G'_{00}$$
$$= G_{00}G'_{00}e^{\beta^*\beta'} \qquad (12\text{-}119)$$

Substitution from Eqs. (8-138) and (8-143) produces, as expected, an F of the form of Eq. (12-104) but with

$$\alpha(t,t') = \frac{C^2}{2\omega}e^{-i\omega(t-t')} \qquad (12\text{-}120)$$

For example, the terms in qq' in Eq. (12-104) come directly from the $\beta^*\beta'$ in the exponential; for this product by Eq. (8-143) is

$$\frac{C^2}{2\omega}\left[\int q(t)e^{i\omega t}\,dt\right]\left[\int q'(t)e^{-i\omega t}\,dt\right] = \frac{C^2}{2\omega}\iint^t [q(t)q'(t')e^{i\omega(t-t')}$$
$$+ q'(t)q(t')e^{-i\omega(t-t')}]\,dt'\,dt \qquad (12\text{-}121)$$

The quantity $a(\nu)$ defined in Eq. (12-109) is therefore [cf. Eq. (5-17) and the Appendix]

$$a(\nu) = \frac{C^2}{2\omega}\int_0^\infty e^{-i\omega t}e^{-i\nu t}\,dt = \frac{C^2}{2\omega}\left[-i\,\text{P.P.}\frac{1}{\omega+\nu} + \pi\delta(\omega+\nu)\right] \qquad (12\text{-}122)$$

so that the real part of $a(\nu)$ is

$$a_R(\nu) = \frac{\pi C^2}{2\omega}\delta(\omega+\nu) \qquad (12\text{-}123)$$

This is zero for positive ν. As expected, we have a "cold environment" as specified in Eq. (12-114).

---

temperature case. For here state n is the initial state with probability proportional to $e^{-\beta E_n}$, so, in view of rule III, the resulting F is obtained by the expression above but with the wave functions $\phi_0(Q_i)\phi_0^*(Q_i')$ replaced by

$$\text{const}\sum_n \phi_n(Q_i)\phi_n^*(Q_i')e^{-\beta E_n}$$

that is, by the density matrix $\rho(Q_i,Q_i')$ worked out in Prob. 10-1. The integrations are again gaussian.

If many independent oscillators of different frequencies are all acting, then by rule IV, their $a_R(\nu)$ functions add; so any cold system (to this gaussian approximation) is equivalent to a continuum of oscillators in their ground state. This follows, since any function $a_R(\nu)$, for negative ν, can be built up of δ functions of the form of Eq. (12-123).

Another interesting example is the interaction with an oscillator at finite temperature. If the temperature is T, the initial state is state n with relative probability $e^{-E_n/kT}$. For our case, the absolute probability is

$$w_n = e^{-n\hbar\omega/kT}(1 - e^{-\hbar\omega/kT}) \tag{12-124}$$

If the initial state were n, the influence functional would be

$$F_n = \sum_m G_{mn}G'^*_{mn} \tag{12-125}$$

instead of the form in Eq. (12-119). Using rule III, we add these with probabilities w_n, so our final F is

$$F = \sum_{m,n} G_{mn}G'^*_{mn}\, e^{-n\hbar\omega/kT}(1 - e^{-\hbar\omega/kT}) \tag{12-126}$$

The sum is difficult to work out directly from Eq. (8-145), but it is

$$F = G_{00}G'_{00}e^{\beta^*\beta'} \exp\left[-\frac{(\beta - \beta')(\beta^* - \beta'^*)}{e^{\hbar\omega/kT} - 1}\right] \tag{12-127}$$

The $a_R(\nu)$ that results from this in place of Eq. (12-123) is

$$a_R(\nu) = \frac{\pi C^2}{2\omega}\left[\frac{e^{\hbar\omega/kT}}{e^{\hbar\omega/kT} - 1}\,\delta(\omega + \nu) + \frac{1}{e^{\hbar\omega/kT} - 1}\,\delta(\omega - \nu)\right] \tag{12-128}$$

and sums of such expressions of many oscillators constitute the environment. Now transitions can go down in energy ($\omega < 0$) or up in energy.

We note that if $\nu > 0$, the first δ function fails, whereas if $\nu < 0$, the second fails, and that indeed

$$a_R(-|\nu|) = e^{\hbar|\nu|/kT}a_R(+|\nu|) \tag{12-129}$$

This definite relation means that in perturbation theory, if $E_n > E_m$,

$$\frac{\text{Probability of a transition per second up } (m \rightarrow n)}{\text{Probability of a transition per second down } (n \rightarrow m)} = e^{-(E_n - E_m)/kT} \tag{12-130}$$

by using Eq. (12-110).

Thus, if the system q occupies states n with relative probabilities $e^{-E_n/kT}$, the net number of up and down transitions will balance out and the system will be in statistical equilibrium for weak perturbation with the environment. This is just what we expect from the laws of statistical equilibrium. Any environment at temperature T producing a quadratic influence functional will have the property of Eq. (12-129).

For an atom as system q in interaction with the electromagnetic field at temperature T as the environment, $a_R(\nu)$ is given by an expression like Eq. (12-128) integrated over all the modes of the field of various frequencies ω. It can be split into the cold environment of Eq. (12-123) plus a noisy external potential:

$$a_R(\nu) = \frac{\pi C^2}{2\omega}\,\delta(\omega+\nu) + \frac{1}{e^{\hbar\omega/kT}-1}\,\frac{\pi C^2}{2\omega}\,[\delta(\omega+\nu)+\delta(\omega-\nu)] \tag{12-131}$$

The first term produces only transitions down in energy and is called spontaneous emission. The second produces transitions up and down with equal ease and is called induced emission, or induced absorption. We say that the transition is induced by an external potential or noise whose mean-square strength at frequency ν varies with temperature as $1/(e^{\hbar\nu/kT}-1)$. This is the way Einstein first discussed the black-body-radiation laws. As we see here, any environment giving a quadratic influence functional at temperature T (we say it is an environment responding linearly) can be treated in the same way. Many people have extended Einstein's argument to other systems, like the voltage fluctuational noise in a resistor at temperature T. The first term measures the rate at which energy is taken out of our system q in a one-way manner. It measures the amount of "dissipation" produced by the environment (e.g., electrical resistance of a metal or radiation resistance of the electromagnetic field). At temperature T we can then say that things behave as if, in addition to the dissipation, there is a noisy signal generated by the environment whose mean square at each frequency is proportional to the dissipation at that frequency and to $(e^{\hbar\nu/kT}-1)^{-1}$. This is called the *dissipation-fluctuation theorem.*

We cannot pursue this subject further here.[1]

---

[1] The subject of influence functionals is discussed in detail by R. P. Feynman and F. L. Vernon, Jr., *Ann. Phys. (N.Y.)*, vol. 24, p. 118, 1963, and W. H. Wells, *Ann. Phys. (N.Y.)*, vol. 12, p. 1, 1961. An application to calculation of mobility of the polaron is in R. Feynman, R. W. Hellwarth, C. K. Iddings, and P. M. Platzmann, *Phys. Rev.*, vol. 127, p. 1004, 1962.

12-10 CONCLUSIONS

In these applications of path integrals to probability theory it is evident that, if the integrands are gaussian, we can make considerable use of the technique. But these problems are precisely those for which other methods, not requiring path integrals, are also available to solve the problem. One may reasonably question the real utility of the path integrals. We can only say that if the problem is not gaussian, it can at least be formulated and studied by using path integrals—and that we might hope that someday, when the techniques of analysis improve, more can be done with it. The only example of a result obtained with path integrals which cannot be obtained in simple manner by more conventional methods is the variational principle discussed in Chap. 11. We hope that further study of these methods may yield more such results.

In the meantime, however, it is worth pointing out that the path integral method does permit a rapid passage from one formulation of a problem to another and often gives a clear or quick suggestion of a relation which can then be more slowly derived in a more ordinary fashion.

With regard to application to quantum mechanics, path integrals suffer most grievously from a serious defect. They do not permit a discussion of spin operators or other such operators in a simple and lucid way. They find their greatest use in systems for which coordinates and their conjugate momenta are adequate. Nevertheless, spin is a simple and vital part of real quantum-mechanical systems. It is a serious limitation that the half-integral spin of the electron does not find a simple and ready representation. It can be handled if the amplitudes and quantities are considered as quaternions instead of ordinary complex numbers, but the lack of commutativity of such numbers is a serious complication.

Nevertheless, many of the results and formulations of path integrals can be reexpressed by another mathematical system, a kind of ordered operator calculus.[1] In this form many of the results of the preceding chapters find an analogous but more general representation (only for the special problems of Chap. 11 is the generalization not known) involving noncommuting variables. For example, in this chapter discussing influence functionals it must have struck the reader that an environment coupled not to the coordinate q but to a noncommuting operator, such as the spin, would be an important and interesting generalization. Such things cannot be conveniently expressed in the path integral formulation but can be easily expressed in the closely related operator calculus.

---

[1] R. P. Feynman, An Operator Calculus Having Applications in Quantum Electrodynamics, *Phys. Rev.*, vol. 84, p. 108, 1951.

An effort to extend the path integral approach beyond its present limits continues to be a worthwhile pursuit; for the greatest value of this technique remains in spite of its limitations, i.e., the assistance which it gives one's intuition in bringing together physical insight and mathematical analysis.

Appendix

Some Useful Definite Integrals

$$\int_{-\infty}^{\infty} e^{ax^2+bx}\,dx = \sqrt{\frac{\pi}{-a}}\,e^{-b^2/4a}$$

$$\int_{-\infty}^{\infty} e^{a(x_1-x)^2}e^{b(x_2-x)^2}\,dx = \sqrt{\frac{-\pi}{a+b}}\exp\left[\frac{ab}{a+b}(x_1-x_2)^2\right]$$

$$\int_0^{\infty}\exp\left(-\frac{a}{x^2}-bx^2\right)dx = \sqrt{\frac{\pi}{4b}}\exp\left(-2\sqrt{ab}\right)$$

$$\int_0^{T}\exp\left(-\frac{a}{T-\tau}-\frac{b}{\tau}\right)\frac{d\tau}{\sqrt{(T-\tau)\tau^3}} = \exp\frac{[-(1/T)(\sqrt{a}+\sqrt{b})^2]}{\sqrt{bT/\pi}}$$

$$\int_0^{T}\exp\left(-\frac{a}{T-\tau}-\frac{b}{\tau}\right)\frac{d\tau}{[\sqrt{(T-\tau)\tau}]^3}$$

$$= \sqrt{\frac{\pi}{T^3}}\,\frac{\sqrt{a}+\sqrt{b}}{\sqrt{ab}}\exp\left[-\frac{1}{T}(\sqrt{a}+\sqrt{b})^2\right]$$

$$\int_{-\infty}^{\infty} e^{-ax^2} = \sqrt{\frac{\pi}{a}}$$

$$\int_0^{\pi/2} e^{-q\sin x}\sin 2x\,dx = \frac{2}{q^2}[(q-1)e^q+1]$$

$$\left.\begin{aligned}&\int_0^{\pi} e^{p\cos x}\sin(p\sin x)\sin ax\,dx\\&\int_0^{\pi} e^{p\cos x}\cos(p\sin x)\cos ax\,dx\end{aligned}\right\} = \frac{\pi p^a}{2a!}$$

$$\int_0^{\infty} e^{-\lambda x^m}x^k\,dx = \frac{1}{m}\lambda^{-(k+1)/m}\Gamma\left(\frac{k+1}{n}\right)$$

Index

Absorption, 153
Acceleration, 174
Action, *viii*, 26–28
Amplitude, 5–9, 13–16, 19–21
  scattering, 142
Annihilation, 232
Approximation of continuity, 218–222
Atomic potential, 134–141
Average, of exponential, 283
  of measured quantity, 112–113
  weighted, of potential, 305

Bethe, H., 256
Black-body Radiation Law, Planck, 294–295
Boltzmann's constant, 268
Born approximation, 128, 131, 169
Born expansion, 128
Bose particles, 16–17, 231, 244, 288
Boson (*see* Bose particles)
Boundary conditions, periodic, 92–94, 213–214
Box, normatizing wave functions in, 89–92
Brownian motion, 165–166, 336–340

Calculus of variations, 26–28
Central-limit theorem, 332
Characteristic function, 325
  of measurement, 106–112
Characteristic functional, 326
Charge density, 237
Chemical constant, 279
Classical path, 59–60
Classical trajectory, 26–28
Cold environment, 350
Communications, theory of, 329
Commutation laws, 176
Commutation between operators, 115–116
Complex conjugate of wave function, 108–109
Compton wavelength, 245
Configuration integral, 278
Conservation of probability, 82–84
Constant field, 63
Continuum, 150–152
Coordinate space compared to momentum space, 101–105
Correlation function, 333–334
Coulomb potential, 137, 239
Coupling coefficient, 70–71
Cross section, 135–139, 152
Crystal, 212–217
  polar, slow electrons in, 310–320
Current density, 237
Cut-off rule, 245–246, 258

d-c voltage, 328–329
Degenerate states, 160
Delta function, 53, 87, 281
Density of levels, 151–152
Density matrix, 273–277
Diffraction, electron, 140
  of light, 11–13
  of particles, 47–57
Dipole approximation, 250
Dirac, Paul A. M., 23, 255
Dirac delta function, 53, 103
Dirac equation, 36
Dispersion, 225
Dissipation-fluctuation theorem, 354

Effective width of diffracting slit, 49–52
Eigenfunctions, 145
  of operator, 115
Eigenvalue for operator, 115
Elastic collisions, 130
Elasticity, 220
Electric dipole, 250
Electromagnetic field, 39, 189–191
Electromagnetic mass correction, 254
Electron spin, 5
Electrons in metal, 293
Emission, 153, 247–250, 260–262
Energy, 28, 46–47, 54, 269
  correction, 253–255
  expansion of kernel, 116–117
  and frequency, 46–47
    in wave functions, 84–86
  levels, 248
  rest, 256
Energy shift, first-order, 159–161
Energy-time transformation, 102–105
Entropy, 272
  of mixing, 291

Environment, cold, 352–353
  time-steady, 350
Environmental system, 343
Exclusive alternatives, 14
Expectation value, 217
Expected value for measurement, 112–115
Extremum, 59–60, 62

Fermi, Enrico, 16
Fermi field, 231
Fermi particles, 16–17, 244, 288–292
Fermion (*see* Fermi particles)
Fine-structure constant, 252
Force, 175, 270
  external, 344
Forced harmonic oscillator, 70–71
Form factor, 138
Four-dimensional symmetry, 237
Fourier series, evaluating path integrals by, 71–73
Fourier transform, 153, 219–220, 226, 314, 325, 335
  as energy-time transformation, 105
Free energy, approximate, 285
Free particle, 42–47, 120
  relativistic, 34–36
Frequency, 54
  and energy, 46–47
  of wave function, 84–86
Fresnel integrals, 49, 55–57
Frohlich, H., 310*n*.
Functional, path integral as, 68–69
Functional derivatives, 170–173

Gaussian function, 49, 52
Gaussian integral, 42–43, 58–62
Gaussian slit for diffraction, 49–54
Geiger counter, 3
Gravitational effects, 246
Ground state, 222–224

Haga, E., 318–319
Hamiltonian, constant, 147
  time-independent, 84
  in vector potential field, 191
Hamiltonian operator, 79–82
Harmonic oscillator, 63, 65, 71–73, 183, 198–234, 276, 286, 351–354
  damped, 336
  forced, 64, 181, 233–234
Harmonic oscillators, independent, 209–211
Heat, 271
Heisenberg, Werner, 9, 161
Helmholtz free energy, 268
Hermite polynomials, 199–200, 203
Hermitian operator, 114
Hermitian property, 82, 86, 188

Independent systems, 67–68
Influence functionals, 343–354
Integral equation, 126, 129, 146
Interaction, 125, 130
  of field and matter, 247–253
  particle and oscillator, 169
  of systems, 66
Interference, 5–8
  pattern, 56
Interfering alternatives, 13–14
Inverse transformation, 111–112

Jacobian, J., 72–73

Kernel, 26, 28–29, 32–36
Kinetic-energy term, 72, 178–179
Kronecker delta, 86, 208

Lagrangian, 26–28, 42
  general quadratic, 58
Lamb, Sir Horace, 253
Lamb shift, 253, 256–260
Laplace, 2
Lattice vibrations, 312
Least action, principle of, 26–28
Lee, T., 318–319
Lifetime, 161
Linear operator, 114
Liquid helium, 287–293

Magnetic analyzer, 45
  field, 64
    analyzer, 100
    and Schroedinger equation, 79–80
Mass, experimental, 257
Matrix element, 145–161
  for transition, 151–161
Maxwell, James C., 23
Maxwell's equations, 23, 230, 237–240
  least-action principle for, 240–242
Mean lifetime, 161
Measurement and probability, 106–112
Measuring equipment and momentum, 100
Meson, 231, 254
Minimum principle, 302–303
Molecule, diatomic, 139
  polyatomic, 140, 203–208
Moment-generating function, 325
Momentum, 8–11, 28, 44–45, 184–189
  amplitude, 96–102
  functional, 187
  measurement, 96
  operator, 116, 185–186
  of photon, 248
  probability, 97–100
  representation, 185
  space, 101–105

Newton's law, 175, 251
Noise, 326–336
  classical, 350
  correlation function of, 334
  gaussian, 329, 331–336
Nondegenerate states, 160
Nonrelativistic approximation, 237
Normal modes, 206–208
  coordinates, 208–212, 216–217
  distribution, 50
Normalization factors, 178
  of wave functions, 89–94
Normalizing constant, 33, 61, 194, 275

Operator calculus, 113
Operator notation, 80–82, 170, 184–189
Operators (*see* Operator notation)
Orthogonal states, 86
Orthonormal functions, 86
Oscillator, field as, 69
  interacting, 205

Partition function, 269–273, 280
  for symmetric statistics, 291
Path, 20–21
Path integral, 34
  double, 341–342
  (*See also* specific problem)
Pauli, Wolfgang, 16
Pekar, S. E., 318*n*.
Periodic boundary conditions, 92–94, 213–214
Periodic paths, 284
Permutation, 288
Perturbation expansion, 120–127
  first-order, 252–253
  time-dependent, 143–161
Perturbed potential, 125
Phase, change of, 45
  of path, 28
Phonons, 222
Photoelectric effect, 153
Photon, 231, 243
Pines, D., 318–319
Planck, M., 243
Planck Black-body Radiation Law, 294–295
Planck's constant, 9
Plesset, M. S., 230*n*.
Poisson distribution, 322
Polar crystal, slow electrons in, 310–320
Polarization, 224
  oscillations of, 312
Polaron, 310
Position measurements, 96
Positrons, 39
Potential, effective, 285
  and momentum representation, 102
Potential-energy term, 72
Potential function, series expansion, 62–63
Power, 153
  mean, 334
  spectrum, 334–336
Pressure, 270

Principal part of integral, 103–104, 156–157
Probability, 2–9, 19, 134
  amplitude, 5–9, 13–16, 19–21
  of function, 323
  and measurement, 108–110
  relative, 43, 51
  of transition, 148–152
Product, of kernels, 38
  of positions, 181
Pulses, random, 322–324

Quadratic actions, 182–184
  approximation, 204
  functional, 184
Quantized field, 69, 229
Quantum electrodynamics, 65, 69, 236–265
Quantum field theory, 229–232

Radiation field, 242
Rayleigh-Ritz method, 307–310
Relativistic particle, 35–36
Relativistic quantum mechanics, 102–103
Relativistic symmetry, 259
Relativistic theory, 38–39, 139
Resonance, 161
Root-mean-square (rms) deviation, 52–53
Rutherford, E., 253
Rutherford cross section, 137

Scalar potentials, 79–80
Scattering, 14–19, 122–127, 129–143, 151
  multiple, 130
  second-order, 147
  x-ray, 138
Schroedinger, Erwin, 22–23
Schroedinger equation, 22–23, 58, 79–81, 129, 195
Schultz, T. D., 319
Schwinger, J., 173, 258
Second-order term, perturbation expansion, 154–158
Separable systems, 66–68
Sound, longitudinal, 225–229
  speed of, 222
Spherical coordinates, 173
Spin, 5, 15–16, 39, 231, 264
Standard deviation, 50–52
State of system, 57
Stationary phase, method of, 132
Statistical problems in quantum mechanics, 340–342, 350–351
Steady states, 84–89
Succession, events occurring in, 36–38
  of velocities, 189
Sum over paths, 29, 31–34
  (*See also* specific problem)
Superconductivity, 293
Symmetry of crystals, 224

Temperature, critical, 292
  finite, 353–354
Thermal equilibrium, 269
Third-order term, perturbation expansion, 158
Time-of-flight as momentum measure, 96–97
Trace, 273
Transformation, of coordinates, 68
  energy-time, 102–105
Transformation functions, 110–112
Transition amplitude, 109, 143–149, 165, 202
  for hamiltonian, 192–195
Transition element, 164–195, 349
  of functional, 168–195
  product of two positions, 181
Transition probability, 248–250
Traveling wave, 222
Trial action, 308

Uncertainty principle, 9–13, 52–54, 85
Unit cell, 224

Vacuum energy, 244–246
Variational method, 296, 303–307
Vector potential, 79–80
  as perturbation, 189–191
Velocity, mean-square, 177–179
Virtual state, 157–158

Voltage, fluctuations of, 329
Volume as normalizing constant, 93–94

Wave, scattered, 142
Wave function, 57–58, 77
  complex conjugate of, 108–109
  expansion for, 127–129
Wave number, 54
  and momentum, 45, 47
Wavelength and momentum, 44–45, 99
Waves, retarded, 251
  running, 242–243
  transverse, 225, 238
Weighted averages, 166
Weisskopf, V. F., 256
White noise, 331

X-rays, scattering of, 138

DATE DUE

GAYLORD

PRINTED IN U.S.A.

MARYGROVE COLLEGE LIBRARY
Quantum mechanics and path int
530.12 F43
3 1927 00030611 5

530.12
F43

Feynman, R. P.
Quantum mechanics and path integrals, by Feynman and Hibbs

DATE	ISSUED TO

530.12
F43

20650